VOLUME 8 NUMBER 3

GENETICS

A PERIODICAL RECORD OF INVESTIGATIONS
BEARING ON HEREDITY AND
VARIATION

MAY, 1923

Published at Menasha, Wisconsin
by the
BROOKLYN BOTANIC GARDEN
Brooklyn, New York, U. S. A.

GENETICS

A Periodical Record of Investigations Bearing on Heredity and Variation

EDITORIAL BOARD

GEORGE H. SHULL, Managing Editor
Princeton University

WILLIAM E. CASTLE
Harvard University

EDWARD M. EAST
Harvard University

EDWIN G. CONKLIN
Princeton University

ROLLINS A. EMERSON
Cornell University

CHARLES B. DAVENPORT
Carnegie Institution of Washington

HERBERT S. JENNINGS
The Johns Hopkins University

BRADLEY M. DAVIS
University of Michigan

THOMAS H. MORGAN
Columbia University

RAYMOND PEARL
The Johns Hopkins University

GENETICS is a bimonthly journal issued in annual volumes of about 600 pages each. It will be sent to subscribers in the United States, the Philippines, Porto Rico, etc., for $6 per annum for the current volume and for $7 per volume for completed volumes. Canadian subscribers should add 25 cents and other foreign subscribers 50 cents for postage. Single copies will be sent to any address postpaid for $1.25 each.

Back volumes 1 to 6 are for sale by WILLIAMS & WILKINS COMPANY, Baltimore, Maryland.

All subscriptions, notices of change of address, remittances, and business correspondence (except concerning volumes 1 to 6) should be sent to GENETICS, Menasha, Wisconsin, U. S. A., or BROOKLYN BOTANIC GARDEN, Brooklyn, New York. All remittances should be made payable to GENETICS.

Application has been made for transfer of second-class entry from Baltimore, Maryland, to Menasha, Wisconsin.

PERMANENCE OF TREE PERFORMANCE IN A CLONAL VARIETY AND A CRITIQUE OF THE THEORY OF BUD MUTATION[1]

KARL SAX AND JOHN W. GOWEN
Maine Agricultural Experiment Station, Orono, Maine

Received December 14, 1922

TABLE OF CONTENTS

INTRODUCTION 179
Material and methods 180
Variation in yields for Ben Davis apple trees 180
Permanence of performance in yield for clonal varieties 184
Differences in tree performance due to soil heterogeneity 190
Ben Davis apple orchard 190
Citrus trees 192
Potatoes 193
Variability in performance of clonal varieties due to variation in stocks 200
Variability in performance of clonal varieties due to bud mutation 202
CONCLUSIONS 207
LITERATURE CITED 210

INTRODUCTION

The investigation presented in this paper is an attempt to analyze the normal variation of yield as it occurs in our common fruit trees, particularly the apple. Our problem and our approach to it may be defined most clearly by considering certain well-known facts concerning variation. In our orchard are brought together a group of homogeneous individuals of the same species and variety, the Ben Davis. If measurements are taken on any character of these individuals, there may be formed from the resulting data a variation curve characteristic for this group and character. Two basic variables are responsible for the position of any particular individual in this curve, environment and heredity. If the character, yield, is measured for another year, another frequency curve may be found. This curve, like the first, has the position of any individual in it defined by the two variables, environment and heredity.

Our problem pertains to these two variation curves. We desire to know for each of them the kind and amount of variation which may take place, and if the defining of the relative position of a given tree for one of the

[1] Papers from the Biological Laboratory, MAINE AGRICULTURAL EXPERIMENT STATION, No. 146.

GENETICS 8: 179 My 1923

Second Printing 1965 / University of Texas Printing Division, Austin

curves will define its relative position for the other curve. In the first question environment and heredity have full play. In the second only that portion of the reaction system, which tends to make the yields of individual trees relatively permanent, can play its part. When compared these two results lead to the defining of two important conclusions; the degree of variation due to uncontrolled environment and the amount of variation due to environment and heredity, which produces permanent influences on yield. The first portion of this paper will be devoted to the analysis of this first problem.

MATERIAL AND METHODS

The data for this study were obtained from Ben Davis Orchard No. 1, at Highmoor Farm, Monmouth, Maine. This orchard came into the possession of the MAINE AGRICULTURAL EXPERIMENT STATION in 1909. The orchard was heavily pruned in 1911 and has been well cared for during the past ten years. The trees are about twenty-eight years old. These trees are set 25 feet apart each way. The orchard contains 1235 trees but on discarding replaced, injured, or diseased trees, or those for which complete records are lacking, there are 882 trees for which we have complete records.

The soil in the orchard is not uniform. The southwest corner of the orchard is high and sandy, while there are low wet areas in the northwest quarter and along the south side. Since 1916 the west half of the orchard has been cultivated while the east half has remained in sod. For the past five years fertilizer experiments have been conducted on rows 5 to 12, inclusive, but there has been practically no effect of such treatment.

Since 1913 annual records have been taken on yield of fruit and growth of each tree. The fruit was weighed in the orchard and all grades were lumped together. The trunk of each tree was measured every fall after the apples were picked. A mark painted on each tree permitted measurements to be made uniformly.

The work of obtaining yield and growth records for the Ben Davis trees was outlined by Dr. FRANK M. SURFACE and until 1919 was under his direction. The credit for the initiation of the work and its careful conduct is due Doctor SURFACE.

The methods used are those in ordinary biometrical use.

VARIATION IN YIELDS FOR BEN DAVIS APPLE TREES

The Ben Davis is a regular bearer, although our orchard has not produced large yields. Table 1 gives the physical constants, mean, standard

deviation and coefficient of variation for the performance of the 882 Ben Davis trees in our orchard for the years 1913 to 1918, inclusive. The unit used is pounds of fruit.

TABLE 1

Physical constants for yields of fruit (pounds) for 882 Ben Davis apple trees.

YEAR	MEAN	STANDARD DEVIATION	COEFFICIENT OF VARIATION
1913	19.9± .4	17.2± .3	86.3±2.2
1914	175.2±2.3	102.3±1.6	58.4±1.2
1915	134.5±1.7	74.3±1.2	55.2±1.1
1916	102.6±1.5	67.4±1.1	65.7±1.4
1917	69.6±1.1	46.1± .7	66.2±1.5
1918	81.1±1.7	72.7±1.2	89.6±2.3
Total 1913–1918	97.1±0.8	84.3±0.6	86.9±0.9

The mean yield of these trees varies between rather wide limits. 1913 was a very low year for the apple-tree yields of this orchard. The succeeding year, 1914, was the high year. The range of variation in mean yield was from 19.9 pounds to 175.2 pounds per tree. The range of the standard deviations is similar to the range found for the mean yields. For 1913 the standard deviation was 17.1 pounds, while in 1914 the standard deviation was 102.3 pounds of fruit. The coefficient of variation is large, ranging from 55.2 for 1915 to 86.9 in 1918, a range of 31.7 percent. Coefficients of variation of this size are among the largest known for data on variation. Thus if other data on reproduction be examined we find that for the number of children per family POWYS has shown the coefficient of variation to be 48.4. For size of litter in mice, for lambs per birth in sheep, annual egg production in the domestic fowl, size of litter in swine, and fecundity in the horse, we find (see PEARL 1919) that the coefficients of variation are, respectively, 37.5 (WELDON), 35.8 (PEARL), 34.2 (PEARL and SURFACE), 27.4 (SURFACE), 24.8 (PEARSON). The coefficients of variation for reproductive ability are some of the largest known for the variables yet measured. It will be noted that the coefficients of variation for the Ben Davis apple yields are even larger than these.

Let us now compare the variability as found in our Ben Davis tree yields with those found for clonal varieties. These data are given in table 2.

TABLE 2

Physical constants of variation in yield of clonal varieties.

KIND OF CLONE	YEAR	MEAN YIELD	STANDARD DEVIATION	COEFFICIENT OF VARIATION	SOURCE OF DATA
Rome apple trees	1914	552.2 ±14.0	154.3 ± 9.9.	27.9 ±1.9	Hedrick and Anthony (1919)
	1915	665.0 ±15.1	166.1 ±10.6	24.9 ±1.7	
	1916	443.1 ±11.9	131.5 ± 8.4	29.6 ±2.0	
	1917	425.0 ±12.8	141.1 ± 9.0	33.2 ±2.3	
	1918	659.5 ±13.7	150.7 ± 9.6	22.8 ±1.5	
Jonathan apple trees	1914	303.9 ± 5.6	125.3 ± 4.0	41.2 ±1.5	Batchelor and Reed (1918)
Washington Navel orange trees	1912	194.8 ± 8.7	102.5 ± 6.1	52.6 ±3.9	Shamel, Scott and Pomeroy (1918 a)
	1913	152.7 ± 6.8	80.5 ± 4.8	52.7 ±3.9	
	1914	210.7 ± 9.8	115.6 ± 6.9	54.8 ±4.1	
	1915	221.0 ±10.2	120.2 ± 7.2	54.4 ±4.1	
Valencia orange	1912	121.7 ± 2.8	43.5 ± 2.0	35.7 ±1.8	Shamel, Scott and Pomeroy (1918 b)
	1913	110.1 ± 2.2	33.3 ± 1.5	30.3 ±1.5	
	1914	115.9 ± 2.3	36.2 ± 1.6	30.4 ±1.5	
	1915	132.2 ± 4.0	61.4 ± 2.8	46.4 ±2.5	
Navel orange (Arlington)	1915	137.6 ± 1.2	54.4 ± .82	39.6 ± .7	Batchelor and Reed (1918)
	1916	186.2 ± 1.7	55.3 ± 1.2	29.7 ± .7	
Valencia orange	1915	246.3 ± 4.3	97.8 ± 3.0	39.7 ±1.4	
Eureka lemon	1915	270.7 ± 2.9	81.4 ± 2.0	30.1 ± .8	
Seedling walnut	1915	99.8 ± 1.9	47.8 ± 1.4	47.9 ±1.6	
	1916	77.6 ± 1.7	41.9 ± 1.2	53.9 ±1.9	
Green Mountain potato	1916	11.30± .19	2.59± .14	26.11±1.28	Whipple (1920)
	1917	26.61± .41	6.25± .29	23.49±1.14	
	1918	17.45± .48	7.47± .34	42.81±2.30	
Rural New Yorker potato	1916	8.80± .16	2.48± .11	28.18±1.39	
	1917	21.98± .22	3.37± .15	15.33± .72	
	1918	19.78± .24	3.77± .17	19.08± .91	

Wide variations in yield of fruit are noted for the different clonal varieties given in table 2. These differences are probably due to differences in soil conditions and age of trees. Similar differences also appear in standard deviation. These are in part to be accounted for by differences in age of the trees. The data on the coefficients of variation for yields are consistent in showing that yields within clonal varieties have a high degree of variability.

The orchard of 55 Rome apple trees at the New York Agricultural Experiment Station is of particular interest in connection with differences in tree productivity. In 1896 Ben Davis trees were selected for their uniformity and planted on uniform soil. These Ben Davis trees were top-worked with buds from a single Rome tree. Data on yield and

treatment of these trees are presented by HEDRICK and ANTHONY (1919). A total of 55 trees are recorded. These trees were divided into 11 plots of five trees each and these plots received different amounts of fertilizers. The effect of the fertilizer treatment has not caused much variability and the variations of individual trees within each plot is much greater than variations between plots. Since the soil was selected for its uniformity and all scions were from a single tree the data are of considerable interest in connection with tree individuality and bud variation. The average yields of the trees are very large as compared with our Ben Davis trees of about the same age. There is considerable variation in individual trees as indicated by the coefficient of variation. The coefficients of variation are much less than those found in our Ben Davis orchard.

SHAMEL, SCOTT and POMEROY (1918 a) give records for four years on 481 Washington Navel orange trees, but the annual detailed performance is recorded for only "64 representative trees from this list." Data on 151 trees for a 6-year period are also given, but complete records are shown for only 35 of these trees. Data for a four year period are given for 105 trees. It is obvious on inspection that the variability of the Washington Navel orange is considerably larger than that for the other orange-tree yields. There is reason to believe that the performance of the Washington Navel orange trees is based on selected data. The original 481 trees were apparently chosen for a study of different strains of the variety. The 64 trees for which complete data are available were said to be "representative trees" from the original group of 481. However, when we apply PEARSON'S χ^2 test for goodness of fit we find that there is not a chance in a million that the two distributions are the same in respect to tree performance. Although the means of the two groups of trees do not differ greatly, we find that the standard deviation of the 64 selected trees is 95.6 ± 5.7, while the standard deviation of the remaining 417 trees is only 58.2 ± 1.4. It is evident that the high- and low-yielding classes are weighted and as a result the variability is greater than would be expected from a true random sample of Washington Navel orange trees. If we make allowance for the weighted data the variability is probably about the same as found in the other orange trees studied.

The variability of the Valencia orange trees is considerably less than that found in the Washington Navel orange, possibly due to the fact that the trees recorded are more nearly a random sample of the entire orchard. The variation is, however, greater than that found in the 55 Rome apple trees.

For the data of BATCHELOR and REED (1918) the variation of the Navel and Valencia orange trees was about the same in amount. This variation compared favorably with that of SHAMEL, SCOTT and POMEROY for the Valencia orange and strengthens the supposition that the data for the Washington Navel group were selected.

If we compare the variation of the orange and the apple it is found that the variation in yield of the Ben Davis was greatest, that of the Jonathan next, and closely approached by the Valencia and Navel oranges, while the variation of the Rome was last. This last is pertinent and is just what one would expect as the orchard was selected for uniform soil, and all trees were top-worked from a single tree. The Rome trees are also very productive, which would reduce the variability, judging from the results of WARING (1920). The relation between yield and variability is also evident from tables 1 and 2. The Ben Davis trees in the 3 unproductive years had a mean yield of about 60 pounds with a variability of 80, while in the 3 productive years the mean yield was 137 pounds with a variability of 60. The Jonathan trees had a mean yield of 304 pounds and a variability of 41, while the Rome trees with an average yield of 540 pounds per tree had a variability of only about 28. The decreased variability of larger-productive individuals might be attributed to a physiological limit of plant growth.

The variation of the potato tuber lines is particularly significant to this study. Here is a clone which can be moved from place to place each year. By changing location, soil differences are smoothed out. Any permanent changes are to be accounted for on the basis of hereditary differences and these differences attributed to some cause like bud mutation. The variability of the potato yields is not significantly different from that found in the Rome apple yields. It is considerably less than that for any of the other clones here presented.

PERMANENCE OF PERFORMANCE IN YIELD FOR CLONAL VARIETIES

In the preceding section it has been shown that yield in clonal varieties is one of the most variable characters known. It now remains to be shown whether or not this variation is of a permanent nature. Stated concretely, are the differences in yield of individual trees consistent from year to year? Surprising as it may seem and in spite of the fact that the advocates of the bud-mutation theory have consistently used the supposed permanence of performance as an argument, our knowledge of the correlation which exists between the yield of one year and the yield of another year for the same tree and within the same clone is, broadly speaking, desultory

and unsatisfactory. The correlation coefficients for the yield of one year with that of another are given in table 3 for our Ben Davis orchard.

TABLE 3

Interannual correlation coefficients and correlation ratios for the yield of 882 Ben Davis apple trees.

	1913	1914	1915	1916	1917	1918
1913		.53 ± .02 .54 ± .02	.53 ± .02 .55 ± .02	.43 ± .02 .44 ± .02	.53 ± .02 .54 ± .02	.26 ± .02 .29 ± .02
1914	.53 ± .02* .54 ± .02†		.86 ± .01 .84 ± .01	.70 ± .01 .71 ± .01	.63 ± .01 .66 ± .01	.36 ± .02 .40 ± .02
1915	.53 ± .02* .55 ± .02†	.86 ± .01 .84 ± .01		.75 ± .01 .74 ± .01	.63 ± .01 .66 ± .01	.40 ± .02 .41 ± .02
1916	.43 ± .02 .44 ± .02	.70 ± .01 .71 ± .01	.75 ± .01 .74 ± .01		.64 ± .01 .65 ± .01	.65 ± .01 .64 ± .01
1917	.53 ± .02 .54 ± .02	.63 ± .01 .66 ± .01	.63 ± .01 .66 ± .01	.64 ± .01 .65 ± .01		.56 ± .02 .57 ± .02
1918	.26 ± .02 .29 ± .02	.36 ± .02 .40 ± .02	.40 ± .02 .41 ± .02	.65 ± .01 .64 ± .01	.56 ± .02 .57 ± .02	

* Correlation coefficients.
† Correlation ratios.

It is evident that there are great differences in the productivity of individual apple trees as indicated by the high values of the coefficients of variation. These differences in productivity are rather consistent from year to year as indicated by the high values for the interannual correlation coefficients. The correlation ratios are not significantly higher than the correlation coefficients, indicating that the regression lines are linear. Since the regressions are linear and the correlation surfaces are uniform the high degree of correlation shows that a true difference in yield exists for the different classes of trees.

The high degree of interannual correlation for yield of apple trees indicates that if a tree is unproductive one year it will be unproductive in succeeding years. By means of a regression equation[2] it is possible to

[2] Where $y = (\text{mean } y - r_{xy}\frac{\sigma_y}{\sigma_x}\text{ mean } x) + r_{xy}\frac{\sigma_y}{\sigma_x}$
y = expected yield
x = actual yield

predict the yield of a given tree in our Ben Davis orchard. Thus, for the group of trees which yielded 20 pounds of fruit in 1914, the average yield of these trees in 1915 was 37 pounds, or for the two years this group of trees was distinctly unproductive. For the group of trees which yielded 200 pounds in 1914 the average yield in 1915 was 150 pounds or this group was distinctly high in production. This condition of affairs is seen to be true for the trees from year to year, the high-yielding trees remaining high, relatively speaking, and the low-yielding trees remaining low-yielding.

The relative permanence of the productivity considered in the light of the permanence of other physiological processes which relate to yield is high for these Ben Davis trees. The range of the correlation coefficients is .26 to .86 with an average of .56. This range compares favorably with that found by GOWEN (1922) for the relation of the milk yields of the 365-day period with subsequent 365-day milk yields for Guernsey (.46 to .81) and Holstein-Friesian cattle (.48 to .78). The wool clip of one season with the wool clip of another had a correlation of .51, according to HILL. The correlation for egg production is shown by HARRIS and BLAKESLEE to be of like size (see GOWEN 1922). All in all, we may conclude that yield of fruit in apple trees agrees with other physiological data on productivity in showing a permanence of yield represented by average correlation of about .56.

As previously pointed out, two sets of factors are acting to produce variability in yield of fruit trees. The first is temporary environmental factors which influence yield. Under this heading should also be included the effect of random sampling. The second is permanent environmental factors and heredity which make permanent differences in yield. These last are measured by the average correlation coefficient of .56. The influence of the first factors on variability of yield is equal to

$$\sigma_x - \sigma_x \sqrt{1-r^2}$$

or in other words their effect is directly proportional to $\sqrt{1-r^2}$. This value is equal to .83 where the correlation coefficient is equal to .56. In other words temporary environmental factors cause 83 percent of the variation in the yield of the Ben Davis trees, whereas the factors which tend to produce permanent differences are the cause of 17 percent of the variability.

Other data on the permanence of yield may now be examined. These data are gathered together in table 4.

TABLE 4

Interannual correlation coefficients for yield in clonal varieties.

KIND OF CLONE	YEARS CORRELATED	COEFFICIENT OF CORRELATION	SOURCE OF DATA
Rome apple	1914–1915	.44 ± .07	HEDRICK and ANTHONY (1919)
	1914–1916	.59 ± .06	
	1914–1917	.19 ± .08	
	1914–1918	.65 ± .05	
	1915–1916	.73 ± .04	
	1915–1917	.70 ± .05	
	1915–1918	.57 ± .06	
	1916–1917	.28 ± .08	
	1916–1918	.59 ± .06	
	1917–1918	.36 ± .08	
Average		.51	
Washington Navel orange	1912–1913	.86 ± .02	SHAMEL, SCOTT and POMEROY (1918 a)
	1912–1914	.71 ± .04	
	1912–1915	.56 ± .06	
	1913–1914	.80 ± .03	
	1913–1915	.61 ± .05	
	1914–1915	.87 ± .02	
Average		.74	
Valencia orange	1912–1913	.63 ± .04	SHAMEL, SCOTT and POMEROY (1918 b)
	1912–1914	.63 ± .04	
	1912–1915	.40 ± .05	
	1913–1914	.45 ± .05	
	1913–1915	.59 ± .04	
	1914–1915	.37 ± .06	
Average		.51	
Green Mountain potatoes	1916–1917	.21 ± .06	WHIPPLE (1920)
	1916–1918	.43 ± .05	
	1917–1918	.52 ± .04	
Average		.39	
Rural New Yorker potatoes	1916–1917	.03 ± .06	
	1916–1918	.04 ± .06	
	1917–1918	.21 ± .06	
Average		.07	

The results for the Rome trees are particularly interesting, for it will be remembered that these trees are scions from a single tree and grafted into Ben Davis stock, the ground being chosen for its uniformity. The correlation coefficient showing the degree of permanence of the variations of these Rome trees is practically equal to that of our Ben Davis orchard.

The permanence in performance of Rome trees is greater than might be expected in view of the low variability.

The Washington Navel orange exhibits much more permanence in its yields than does the Valencia orange or either sets of apple data. It will be remembered that the Washington Navel orange data was shown to be highly selected. As there is no means of knowing what the limits of this selection were, it is probably best not to consider these data as representative.

The data on the Valencia orange were apparently a random sample of the whole population. Here the degree of permanence of yield was approximately the same as that of the apple data. In these three clonal varieties there is about 17 percent of the variation which is permanent.

Studies on bud variation in the Lisbon lemon and Eureka lemon have also been conducted by SHAMEL, SCOTT, POMEROY and DYER (1920). The complete records of only 25 Lisbon lemon trees and 24 Eureka lemon trees are presented. The interannual correlations for yield of 25 Lisbon lemon trees are as follows: For 1913–1914, $r = .60 \pm .08$; for 1913–1915, $r = .55 \pm .09$; and for 1914–1915, $r = .75 + .06$. Here again the results do not differ greatly from those found for the larger series of data.

The permanence of variations in yield within a clonal variety of potatoes has the same basis as permanence of tree performance. The potatoes have the advantage, however, in that the effect of environment can be largely eliminated. In tuber lines of potatoes soil differences are smoothed out by changing the location of plots each year. If certain lines remain consistently high or low, year after year, and in different plots of soil, we can assume that the differences are due to inherent characters of the tuber line. We must assume, of course, that the tuber lines are free from disease. The performance of a fruit tree on the other hand may be consistently low or high for a series of years due to the constant soil differences or to the effect of the stock on the scion. In the experiments of WHIPPLE (1920) the seed pieces were the same weight and all plots received the same treatment. The permanence of performance of individual tuber units as indicated by interannual correlations for yield is shown in table 4. The average of the interannual correlations for yield is $r = .39$ for the Green Mountain variety, while there is practically no interannual correlation for yield in the tuber lines of the Rural New Yorker. In both varieties it will be noted that the interannual correlations for yield are greater for 1917–1918 than for 1916–1917.

According to WHIPPLE the Green Mountain variety has degenerate tendencies. "Degeneration(?)" is indicated by the presence of "curly-dwarf" plants. These curly-dwarfs may appear suddenly. FOLSOM, SCHULTZ and others (1919–1920) have recently shown that so-called degeneration is frequently, if not always, associated with the presence of infectious diseases. Potato mosaic is a very common disease of potatoes and according to MELHUS (1917) mosaic-infected plants from Maine, when planted in Iowa, developed curly-dwarf. According to SCHULTZ, FOLSOM, HILDEBRANDT and HAWKINS (1919) mosaic-infected potatoes from Maine may show no mottling when grown in Colorado. The similarity of mosaic and curly-dwarf suggest that they are due to the same or similar causes. BARRUS (1918) states that varieties of Green Mountain potatoes are most seriously affected with mosaic, while the Rural New Yorker type is rarely seriously or even slightly affected. SCHULTZ and FOLSOM (1920) have shown that potato mosaic and curly-dwarf are transmitted from infected to healthy stock by aphids.[3] The work of FOLSOM and SCHULTZ make the interpretation of WHIPPLE'S results fairly clear. The Green Mountains were evidently infected with curly-dwarf at the beginning of the experiment and in successive years the diseased tuber lines maintained a low production. The sudden appearance of degenerate types in healthy lines was probably due to infection from neighboring diseased plants. The Rural New Yorkers were found to be "quite stable," which would be expected because they are little affected by the forms of virus diseases normally found in Green Mountains. We may conclude then that the clonal variety of Green Mountains was rather variable and that variations in yield were relatively permanent because of the presence of diseased lines. In the Rural New Yorkers, little affected by disease, we find practically no permanence of variations in performance.

EAST (1910) finds no permanence in high and low selections within a clonal variety of potatoes. STUART (1915) has selected high- and low-yielding lines of potatoes but attributes the differences in yield to disease.

GREEN and WAID (1906) obtained greater yields from tubers of high-yielding parents than from tubers of low-yielding parents. However, as EAST (1910) points out, the size of the seed piece and disease might account for all differences in performance of selected lines.

It is evident that there is considerable variability in the performance of individual trees within a clonal variety. Such differences are relatively

[3] The writers are indebted to Dr. D. FOLSOM of the MAINE AGRICULTURAL EXPERIMENT STATION, for suggestions in regard to potato diseases, but the writers alone are responsible for the conclusions drawn.

permanent as indicated by the high values of the interannual correlations in all cases. When we make allowance for weighted data the consistent differences in performance of citrus trees are no greater than in the Ben Davis apple trees. The Rome apple trees are not as variable as the Ben Davis apple trees or orange trees although the differences in performance are just as permanent. A study of tuber lines of potatoes indicates that when only healthy plants are considered there are no consistent differences in productivity within a clonal variety. All clonal varieties of fruit trees clearly show permanent individual variations in yield of fruit. The causes of these differences in performance of apple and citrus trees, may in the authors' opinions, be attributed primarily to the following known factors: (1) Soil heterogeneity; (2) effect of stocks; and (3) bud variation. The effects of these factors in causing variability within clonal varieties will be considered in the above order.

DIFFERENCES IN TREE PERFORMANCE DUE TO SOIL HETEROGENEITY

Ben Davis apple orchard

By grouping the trees in our Ben Davis orchard in a four-by-four-fold manner, sixteen trees to the block, and considering them as such we should largely eliminate tree individuality, due to the effect of stock and scion, in causing variability in yield. The influence of the remaining variable, soil heterogeneity, can then be analyzed. The Ben Davis orchard No. 1 contains 37 rows. The number of trees per row varies from 37 at the north end to 30 at the south end. If we consider only the trees in rows 1 to 36, inclusive, and trees 1 to 32, inclusive, we have a block of trees which can be grouped in a 4-by-4-fold manner into 72 plots of 16 trees each. The mean tree yields and variability of the plots are recorded in table 5.

TABLE 5

Annual yields of 72 plots of Ben Davis apple trees. 822 trees.

YEAR	MEAN	STANDARD DEVIATION	COEFFICIENT OF VARIATION
1914	176.67±5.67	71.34±4.01	40.38±2.61
1915	137.08±4.14	52.04±2.93	37.96±2.42
1916	100.42±3.35	42.20±2.37	43.02±2.83
1917	67.50±2.51	31.52±1.77	46.70±3.15
1918	73.75±3.90	49.10±2.76	66.58±5.12

The mean tree yields of the plots are in most cases practically the same as for the average tree yields of the orchard as a whole. The coefficient of variation, although high, is considerably lower than in the case of individual trees. We may conclude then that although there is considerable variation in plots of apple trees due to general soil differences, yet the greater variability of individual trees indicates that small pockets of soil, differences in stocks, or inherent differences in individual trees, either singly or collectively, are partly responsible for difference in tree performance.

The differences in performance of plots of apple trees are about as permanent from year to year as differences in performance of individual trees. The interannual correlations for yield of the 72 plots are presented in table 6.

TABLE 6

Interannual correlations for yield of 72 plots of Ben Davis apple trees.

	1914	1915	1916	1917	1918
1914		.86± .02	.62± .05	.50± .06	.06± .08
1915	.86± .02		.67± .04	.47± .06	.14± .08
1916	.62± .05	.67± .04		.43± .07	.59± .05
1917	.50± .06	.47± .06	.43± .07		.33± .07
1918	.06± .08	.14± .08	.59± .05	.33± .07	

The average interannual correlation between yields of plots is nearly as great ($r = .47$) as the interannual correlation for yield of the 882 individual trees ($r = .56$). Some disturbing influence in 1918, lowered the interannual correlations of plots considerably. In general, however, there is a high correlation between the yields of plots in successive years. The high-yielding plots are permanently low-yielders to about the same extent that variations in individual trees are permanent.

It is evident that the soil is not uniform in Ben Davis orchard No. 1, and that soil differences are permanent. If we apply the measure of soil heterogeneity proposed by HARRIS (1915), we find the soil to be highly heterogeneous when measured by yield and growth of apple trees. The orchard was divided into 72 plots of 16 trees each as the ultimate plots. The ultimate plots were grouped in a 2-by-3-fold manner making 12 combination plots. The degree of soil heterogeneity for the years 1914 to 1918, inclusive, is shown in table 7.

TABLE 7

Ben Davis apple orchard. Soil heterogeneity.

YEAR	MEASURED BY YIELD OF 16 TREE PLOTS AS THE ULTIMATE UNIT	MEASURED BY YIELD OF INDIVIDUAL TREES AS THE ULTIMATE UNIT	MEASURED BY GROWTH OF 16-TREE PLOTS AS THE ULTIMATE UNIT
1914	.57 ± .05	.39 ± .06	.43 ± .07
1915	.44 ± .06	.41 ± .06	.25 ± .08
1916	.39 ± .06	.37 ± .06	.71 ± .04
1917	.47 ± .06	.39 ± .06	.58 ± .06
1918	.48 ± .06	.41 ± .06	.71 ± .04

The average degree of soil heterogeneity as measured by yield is $r = .47$ and when measured by growth, $r = .54$. The soil heterogeneity determined by using the average yield of the 16-tree plots as the ultimate unit gives a higher constant than when the ultimate unit is a single tree, because of the small numbers of trees in some plots, particularly in the most heterogeneous portion of the orchard. By using the method as originally proposed by HARRIS (1913) for cases where individuals in certain plots are missing we are able to determine the soil heterogeneity for the individual tree as the ultimate unit and combining in a 4-by-4-fold manner to form the 72 combination plots. The soil heterogeneity obtained by using individual trees as the ultimate unit is rather high ($r = .40$). Much of the variation found in the plots and trees of Ben Davis orchard No. 1 is evidently due to soil heterogeneity.

Citrus trees

SHAMEL, SCOTT and POMEROY (1918) believe that differences in citrus tree performance as noted by them are not due primarily to the effect of soil. Unfortunately the trees selected are generally in small groups. The Washington Navel orange trees studied by SHAMEL and others were located in two adjacent groves. In view of the belief that soil plays no important part in tree differences, it is surprising to find the different strains in isolated groups. Of the 481 trees recorded only 1.7 percent are of the Golden Nugget strain and yet half of the Golden Nugget trees are found in a block of 15 trees (rows 15 to 17, trees 19 to 23, inclusive). In this block of trees 26.6 percent are of the Golden Nugget strain. In the same way 73.2 percent of the trees in a block of 56 trees are of the Thompson strain, although the percentage of Thompson strain in the entire group of 481 trees is only 10.6 percent. The trees of the Unproductive strain are in several groups throughout the orchard. In one

group of 27 trees 51.8 percent are the Unproductive strain although in the whole group only 7.3 percent of all trees recorded are the Unproductive strain. The different plots were, of course, selected because they contained certain strains of trees, but the fact that different strains can be found in groups clearly indicates that soil is responsible for much of the difference in tree performance. A similar grouping of unproductive trees is found in the Valencia orange and in the lemon trees recorded by SHAMEL and others. An examination of table 2, United States Department of Agriculture Bulletin 815, shows the striking relation between location, type, and yields of the various trees. Such distributions of tree types and differences in performance may be largely attributed to soil heterogeneity. At least the effect of soil is an important factor in tree performance.

Potatoes

If variations in performance of individuals are inherent and permanent, we would expect that in general high-yielders would remain relatively high-yielders and low-yielders would remain relatively low-yielders under different environmental conditions. The fact that a fruit tree is a consistently high- or low-yielder may simply indicate that it is located in good or poor soil or that it is grafted on a good or poor stock. It is impossible to move mature trees to new plots each year, but we already have some data along the same line with potatoes, which is of interest in this connection. The constants obtained from the data on WHIPPLE'S (1920) Rural New Yorker potatoes show that there are no permanent variations in performance of tuber lines. This variety was selected because it is apparently relatively free from disease and only healthy plants should be considered. The plots of potatoes were not permanent, but were presumably changed according to a system of rotation. According to WHIPPLE the soil was uniform and all plots received the same treatment. Although the average variability for the three years, 1916 to 18, is rather high, there is practically no permanence in performance of individual tuber lines. The high-yielders of one year are not consistently high-yielders for a series of years when the lines are grown on different plots. It is evident that all of the differences in production found in individual tuber lines are due to the effects of environmental conditions.

The evidence thus far presented to show that differences in tree performance within a clonal variety are due to bud variations is based on the fact that apple and citrus trees under apparently identical environmental conditions vary greatly in productivity and that some of this variability

is permanent. Of interest in this connection are some results of HARRIS and SCOFIELD (1920). The degree of permanence in yields of plots is indicated by the interannual correlations for yield. When a crop is grown on the same plots for several years in succession, it was found that certain plots were consistently high-yielders while other plots were consistently low-yielders. The crops considered were millet grown in India, corn grown at the ILLINOIS AGRICULTURAL EXPERIMENT STATION, hops grown in California, and alfalfa grown in Montana. The most extensive records analysed were from the Field Station of the OFFICE OF WESTERN IRRIGATION AGRICULTURE, at Huntley, Montana. At this Station crops were grown on 46 permanent plots since 1911. Correlations were obtained between the same crops and different crops in the different years. The coefficients obtained indicate that even on soil of apparent uniformity, the plots which yield higher one year will yield higher for a series of years even when planted with different crops. The correlations for yield of plots of alfalfa are especially interesting as the same crop is grown on the plots for three successive years. The correlation for total yield of plots of alfalfa for various years are as follows: 1912–1913, $r = .250 \pm .093$; 1913–1914, $r = .886 \pm .021$; and for 1912–1914, $r = .429 \pm .081$. For purposes of comparison these constants are summarized with the data on clonal varieties of fruit trees and potatoes. The constants for the data on corn were taken only for 1896 and 1897, since in 1895 the crop was practically a failure and not representative for interannual comparisons.

On examining table 8, we find that permanence in performance is not limited to fruit trees. Permanent plots of millet, alfalfa, and corn, when grown for a series of years show just as much permanence of differences in performance as do individual plots of soil containing a single apple or orange tree. In the case of plots of millet, corn, or alfalfa the permanence of differences in yield cannot be ascribed to inherent variations of individual plants. Variations in behavior of individual grain or forage plants would be smoothed out owing to the large numbers on each plot and any differences in performance of individual plots must be due to soil differences. The fact that differences in performance of individual plots, regardless of the crop planted, are relatively permanent, clearly indicates that permanent differences in yields of fruit trees of a clonal variety may be due to soil heterogeneity and not to inherent differences in individual trees.

If the differences in performance of individual trees are due to the effects of soil we would expect to find some correlation between the varia-

bility of tree performance and soil heterogeneity, i.e., if the coefficient of variability of the trees of an orchard is large, then we would expect a high degree of soil heterogeneity. The table on page 196 shows the relation between variability in yields of orange, lemon, walnut and apple trees and the degree of soil heterogeneity in each case.

In all cases the individual tree was used as the ultimate unit and combined in a two-by-two-fold manner to form the combination plots. The Valencia orange trees recorded by SHAMEL, SCOTT and POMEROY (1918) include rows 85 to 90, trees 1 to 10, inclusive. The Eureka lemon trees (block 18) include rows 73 to 75, inclusive, as given by SHAMEL, SCOTT,

TABLE 8

Average variability and average interannual correlation for yield of various crops.

CROP	LOCATION	YEARS (INCLUSIVE)	NUMBER OF TREES OR PLOTS	ACREAGE PER TREE OR PLOT	AVERAGE COEFFICIENT OF VARIATION	AVERAGE INTERANNUAL CORRELATION
Washington Navel orange.	California	1912–1915	63	.....	53.66	.74
Valencia orange	California	1912–1915	105	.....	35.72	.51
Lisbon lemon	California	1912–1915	25	.....	36.44	.63
Rome apple trees.....	New York	1914–1918	55	.037	27.73	.51
Ben Davis apple trees.....	Maine	1914–1918	95	.014	46.02	.65
Ben Davis apple trees	Maine	1913–1918	882	.014	70.22	.56
Plots of Ben Davis apple trees.....	Maine	1914–1918	72	.224	46.93	.47
Millet plots (permanent)..	India	1905–1907	105	.....	33.6	.74
Plots of alfalfa (permanent)	Montana	1912–1914	46	.170	13.39	.52
Plots of corn (permanent)..	Illinois	1896–1897	120	.100	10.2	.82
Plots of potatoes. Rural New Yorker (rotation)	Montana	1916–1918	108	.002	20.86	.09

POMEROY and DYER (1920b table 7, page 29). The Lisbon lemon trees, (block 17) include rows 26 to 35, trees 8 to 17, inclusive, as given by the same authors (1920a, table 2, page 16).

In some cases small blocks of trees in large orchards were selected because of their apparent uniformity. These blocks were selected as follows: Block 2, rows 1 to 10, trees 41 to 50, inclusive, in the Navel orange orchard (block 1); block 3, rows 1 to 12, trees 1 to 4, inclusive, in block 1; block 4, rows 30 to 38, trees 1 to 6, inclusive, in block 1; block 7, rows 1 to 10, trees 7 to 12, inclusive, in the Valencia orange orchard (block 6); block 11, rows 5 to 10, trees 1 to 10, inclusive, in block

10; block 13, rows 37 to 44, trees 1 to 8, inclusive, in block 12; and block 15, rows 2 to 11, trees 18 to 28, inclusive, in the Ben Davis orchard (block 14).

In the Washington Navel and Valencia orange groves recorded by BATCHELOR and REED (1918) there is a significant decrease in variability of the selected plots associated with a significant decrease in soil heterogeneity as compared with the entire orchards. In the Valencia orange

TABLE 9

Variability in yield of fruit trees and soil heterogeneity.

CROP	YEAR	ACREAGE PER TREE	NUMBER OF TREES	COEFFICIENT OF VARIATION	SOIL HETEROGENEITY
1 Navel orange (Arlington)*	1915	.011	1000	39.55± .68	.517±.016
2 Navel orange (Arlington)†	1915	.011	100	25.39±1.29	.245±.063
3 Navel orange (Arlington)†	1915	.011	48	29.26±2.18	.098±.096
4 Navel orange (Arlington)†	1915	.011	54	21.35±1.45	.090±.091
5 Navel orange (Antelope Heights)*	1916	.011	495	29.72± .69	.375±.026
6 Valencia orange*	1915	.011	240	39.72±1.40	.306±.039
7 Valencia orange	1915	.011	60	31.64±2.13	.169±.085
8 Valencia orange‡	1913	.....	60	23.48±1.51	.313±.078
9 Eureka lemon*	1915	.013	364	30.06± .81	.448±.028
10 Seedling walnut*	1915–1916	.057	280	46.41±1.58	.232±.038
11 Seedling walnut†	1915–1916	.057	60	47.51±3.52	−.033±.087
12 Jonathan apple*	1914	.011	224	41.23±1.52	.214±.043
13 Jonathan apple†	1914	.011	64	35.64±2.38	−.068±.084
14 Ben Davis apple	1914	.014	822	58.40±1.22	.396±.020
15 Ben Davis apple	1914	.014	95	42.78±2.45	−.069±.069
16 Rome apple**	1918	.037	60	26.16±1.79	.002±.090
17 Lisbon lemon‡	1913–1917	.....	100	22.62±1.13	.216±.064
18 Eureka lemon‡	1911–1917	.....	80	24.12±1.36	.341±.067

* Data and Coefficient Variation from BATCHELOR and REED (1918). Soil heterogeneity calculated by HARRIS (1920).

† Data from BATCHELOR and REED (1918).

** Data from HEDRICK and ANTHONY (1919).

‡ Data from SHAMEL, POMEROY AND DYER (1920).

orchard and Eureka lemon trees recorded by SHAMEL and his co-workers, a relatively low variability of tree performance is associated with a relatively high degree of soil heterogeneity, indicating that much of the variability may be due to the effect of soil. The high degree of variability of the walnut seedlings is not reduced by selecting a block of 60 trees in soil that is found to be homogeneous. We would expect that most of the variability of the seedling walnuts would be due to tree individuality because of their

probable heterozygosity. In the case of the Jonathan apple trees the decrease in variability and the decrease in soil heterogeneity as the result of selecting an apparently uniform plot are hardly significant in either case. In the Ben Davis orchard there is considerable reduction in variability associated with the complete elimination of soil heterogeneity. Most interesting is the rather high variability of the Rome apple trees of the NEW YORK AGRICULTURAL EXPERIMENT STATION, which were propagated from a single tree and are on soil that is, in general, uniform. It is evident that in many cases tree variability is independent of general soil differences.

The consistent differences in yield on Ben Davis apple trees independent of general soil differences, can be shown by eliminating the effects of soil differences. Since the number of trees in each of the 72 plots varies it was considered necessary to hold this factor constant also. Accordingly all possible correlations were made between (1) the tree yield one year; (2) the tree yield of the following year; (3) the plot yield of one year; (4) the plot yield of the following year; and (5) the number of trees per plot. These correlations are shown in table 10.

For instance the correlation between the yield of trees in 1914 and 1915 is .82, and the correlation between yield of trees in 1914 and number of trees per plot is $-.04$. It is apparent that the number of trees per plot does not have much effect on consistent difference in tree yields, but because of possible differential death rate or increased space around isolated trees it was considered necessary to determine the effect of this factor.

Now if plot yields and the number of trees per plot are held constant it is possible to measure the permanence of differences in tree yields independent of the effect of general soil differences. In other words the permanence of differences in yield of trees due to differences of stock, inherent differences in individual trees, or perhaps local soil pockets, can be measured.

The interannual correlations between individual tree yields with plot yields and number of trees per plot held constant are shown in the last column of table 10.[4] These constants are not significantly less than the interannual correlations without soil corrections. These results indicate that individual-tree variability is independent of plot variability. In

[4] The partial correlations were obtained by the formula

$$r_{12\cdot 345}=\frac{r_{12\cdot 34}-r_{15\cdot 34}\,r_{25\cdot 34}}{\sqrt{1-r_{15\cdot 34}^2}\ \sqrt{1-r_{25\cdot 34}^2}}$$

TABLE 10

Interannual correlations for yield of trees and yield of plots and correlations between yield of trees or plots and number of trees per plot. Also interannual correlations of tree yields with other factors constant. 1 = 1st year of tree yield; 2 = 2nd year of tree yield; 3 = 1st year of plot yield; 4 = 2nd year of plot yield; 5 = trees per plot.

YEARS	FACTORS CORRELATED AND CORRELATION COEFFICIENTS										
	1, 2	1, 3	1, 4	1, 5	2, 3	2, 4	2, 5	3, 4	3, 5	4, 5	$r_{12\cdot 345}$
1914–1915	.82 ± .01	.66 ± .05	.59 ± .05	−.04 ± .02	.57 ± .06	.65 ± .05	−.03 ± .02	.86 ± .02	−.07 ± .08	−.09 ± .08	.75
1914–1916	.67 ± .01	.66 ± .05	.42 ± .06	−.04 ± .02	.40 ± .07	.63 ± .05	.25 ± .02	.62 ± .05	−.07 ± .08	.34 ± .07	.68
1914–1917	.60 ± .01	.66 ± .05	.35 ± .07	−.04 ± .02	.34 ± .07	.63 ± .05	.17 ± .02	.49 ± .06	−.07 ± .08	.17 ± .08	.61
1914–1918	.32 ± .02	.66 ± .05	.07 ± .08	−.04 ± .02	.07 ± .08	.67 ± .05	.37 ± .02	.06 ± .08	−.07 ± .08	.57 ± .05	.14
1915–1916	.73 ± .01	.65 ± .05	.44 ± .06	−.03 ± .02	.44 ± .07	.63 ± .05	.25 ± .02	.67 ± .05	−.09 ± .08	.34 ± .07	.74
1915–1917	.56 ± .02	.65 ± .05	.31 ± .07	−.03 ± .02	.31 ± .07	.63 ± .05	.17 ± .02	.47 ± .06	−.09 ± .08	.17 ± .08	.60
1915–1918	.38 ± .02	.65 ± .05	.11 ± .08	−.03 ± .02	.12 ± .08	.67 ± .05	.37 ± .02	.14 ± .08	−.09 ± .08	.57 ± .06	.51
1916–1917	.57 ± .02	.63 ± .05	.32 ± .07	.25 ± .02	.36 ± .07	.63 ± .05	.17 ± .02	.43 ± .07	.34 ± .07	.17 ± .07	.53
1916–1918	.58 ± .01	.63 ± .05	.39 ± .07	.25 ± .02	.40 ± .07	.67 ± .05	.37 ± .02	.59 ± .05	.34 ± .07	.57 ± .06	.59
1917–1918	.50 ± .02	.63 ± .05	.31 ± .07	.17 ± .02	.30 ± .07	.67 ± .05	.37 ± .02	.33 ± .07	.17 ± .08	.57 ± .06	.43

other words tree variability is superimposed on a relatively high degree of soil variability.

The analysis of the available data on clonal varieties shows clearly that soil heterogeneity plays a considerable part in influencing productivity and in establishing differences in the yields. A point not so obvious, however, is the fact that soil heterogeneity has no effect on the *permanence* of variability in tree productivity. For example, the coefficient of variability was reduced from 58.4 to 42.8 by selecting a block of 95 trees in the Ben Davis orchard and the soil heterogeneity was reduced from .40 to .07, *but the permanence of the variability of the 95 trees was not reduced at all.* In fact the average interannual correlation for the 882 trees was .56 while for the block of 95 selected trees the average interannual correlation was .65. This same point is more conclusively shown in table 10. When the influence of soil is eliminated and all differences except inherent differences in scions or stocks, or possibly those due to local soil pockets, are held constant, the permanence of variability is not reduced. This is evident from the fact that the average value or r_{12} is .57 while the average value of $r_{12 \cdot 345}$ is .56 (table 10).

The effect of the various factors on variability can be determined by holding all but one factor constant. Thus the standard deviation of tree yields is 45 for our Ben Davis trees when factors causing permanence in tree variability (stocks, scions, or unknown factors), the effects of soil, and the number of trees per plot, are held constant. (Mean value of constants for 1914–1918 was used.) When we consider that the mean standard deviation of these trees is 72 when all of these factors causing variability are included, it is evident that only 38 percent of the variability can be controlled by selecting uniform stocks, scions and soil, and that 62 percent of the variability is due to causes such as weather, pollination, and unknown factors. When all factors considered are controlled the standard deviation of the yield is 45 and when all factors except soil differences are controlled the standard deviation is 58. The difference between these values is 13, the standard deviation of tree yields due to the effects of heterogeneous soil. In other words 18 percent of the tree variability is due to soil differences. In a similar way it can be shown that only about 15 percent of the tree variability is due to factors which cause permanence in differences of tree performance, i.e., stocks, bud mutation, and unknown factors.

The only effect of heterogeneous soil or other environmental factors is to increase the variability of tree productivity (assuming sufficient numbers for the effect of random sampling), but there is no effect of these

factors on the *permanence per se* of variations in productivity. The effect of one of the factors which may cause permanence in differences in yield of fruit trees will be considered in the next section.

VARIABILITY IN PERFORMANCE OF CLONAL VARIETIES DUE TO VARIATION IN STOCKS

The preceding discussion shows that soil differences do not account for the permanence of variations in yield. It has also been shown that Washington Navel orange, Valencia orange, seedling walnut, Jonathan apple, Ben Davis apple and Rome apple trees may be variable in performance even on soil that is, in general, homogeneous. In the Ben Davis apple orchard permanent differences in the yields are found even when general soil differences are eliminated. It is evident that the great variability of clonal varieties of fruit trees is not entirely due to soil heterogeneity, but must be due to some extent to tree individuality.

Tree individuality may depend either on bud variations propagated by buds or scions, or on the variations found in the stocks. Bud variations involving even qualitative characters in the apple are very rare and it has not been shown that differences in performance are due to bud variation in any clonal variety. The fact that clonal varieties can be maintained practically unchanged for many generations through almost infinite numbers of cell divisions, and under diverse conditions, suggests that, in apple trees at least, the occurrence of bud variations is not a very large factor in variability of fruit trees within a clonal variety. On the other hand little attention has been paid to the characters of the stocks used for grafting. In most cases the stocks are seedlings from horticultural varieties. These varieties are undoubtedly heterozygous, and since most varieties of apples are self-sterile, they are probably also cross-pollinated. The seeds from such a variety are highly heterozygous and give rise to seedlings which are extremely variable in growth and vigor as well as for morphological characters. When these seedlings are used as root stocks it is only logical to suppose that the scions are greatly influenced in regard to vigor, growth and productivity.

It is well known that different varieties of stock cause large variations in the growth of the top-worked bud or scion. The standard stocks permit a normal growth of the scion while dwarf stocks check the growth of the scion and a dwarfed tree is produced. Stocks which produce a weak small tree would also produce decreased yields. The stocks commonly used in commercial nursery work are seedlings of the "French Crab."

According to HEDRICK (1915), "What we in America call standard stocks, in England more commonly called free stocks, known also in both countries as Crab and French Crab stocks, are seedlings of the common apple, *Pyrus malus*." Crab seeds are rarely planted for stocks. Since the French Crab seed is selected not from a single variety, but from wild trees and unknown varieties, great variability of the resulting seedlings would be expected. When apple seed is taken from cider-mill refuse for growing stocks the variability of the resulting seedlings must be very great as HEDRICK has suggested. Not only are seedling stocks variable because of random selection of seed from different types, varieties and even species of apples, but seeds from a single variety or even a single tree will result in extremely variable seedlings in most cases. Most varieties of apples are highly heterozygous as is indicated by the seedlings resulting from controlled pollinations. We have at our experimental farm over 1400 seedlings, most of them from the leading apple varieties grown in New England. These seedlings are extremely variable, most of them resembling more or less closely the types of trees which grow wild and are known as "naturals." A group of 586 trees are from seeds planted in 1911. The circumference of these trees varies from 2 to 18 cm with a mean of $8.92 \pm .08$. The standard deviation is $2.86 \pm .06$ and the coefficient of variability is $32.07 \pm .69$. Although these seedling trees are extremely variable they are undoubtedly at least as uniform as seedlings grown for stocks.

In our "stock and scion" orchard we have over 400 trees of ten varieties. When these trees were taken from the nursery they varied considerably in size. If these differences in size were due to differences in the soil of the nursery we would not expect these differences to be permanent when the trees were transplanted in the orchard. The random planting of large and small trees should smooth out differences in growth if soil is the only factor involved. If, however, the variation in size of nursery trees is due to the effect of stocks or scions, then differences in size will persist in the trees when set in the orchard. The buds of each variety were selected from a single tree in each case, thus largely eliminating any possible difference in growth of scion due to bud mutation. The growth rate of different varieties on the two kinds of stocks was not found to be the primary cause of correlation between size of nursery tree and the same tree in the orchard. We have found that the trees which are small when set in the orchard are also the small trees in succeeding years. The correlation between circumference in 1916, two

years after the trees were set, and the circumference in 1921, was found to be very high ($r = .68$). We may conclude then that small trees in the nursery, in general result in small trees in the orchard, due to a considerable extent to the effect of the stock.

The variability in performance of citrus trees is very probably largely due to soil heterogeneity and to variations in stocks used. WEBBER (1920) has recently called attention to the necessity of selecting more uniform stocks for citrus fruits. He shows that the seedlings commonly used are variable in vigor and size as well as in qualitative characters. Sweet seedling stock was top-budded with buds from recorded trees. Small and large trees in the nursery retained this size difference when planted in the orchard. WEBBER concludes that the stocks used must cause the permanent variability in size found in the trees. In order to obtain more uniform stocks WEBBER (1920) suggests that

> "We must no longer grow merely sour stock or sweet stock and the like. The process must be carried farther and good stock varieties of sour orange and sweet orange must be discovered and named as stock varieties and every nurseryman should then use seeds from these varieties known to produce good stock seedlings."

However, since all citrus trees are probably heterozygous and in addition are cross-pollinated the seeds from a single tree would undoubtedly produce various types of seedlings. If clonal varieties of desirable stocks could be established and used the results would be more uniform and certain.

The data, thus far presented, have shown that soil heterogeneity causes permanent, marked variations in tree performance. The permanence of tree performance is accentuated by the stock. The character of trees grown on dwarf stocks, the variability of seedling stocks, and the permanence of these variations in the budded tree establishes this point. It now remains to be seen how much the permanence of tree performance is dependent on the character of the bud or scion.

VARIABILITY IN PERFORMANCE OF CLONAL VARIETIES DUE TO BUD MUTATION

The occurrence of bud variations involving characters such as color, shape and size is well established. In most of our horticultural fruit varieties bud variations are rare, but in citrus fruits they are apparently more numerous and are often of considerable economic importance. There is little critical evidence, however, that bud variations occur which involve productivity. The results of the experiments on bud selection for produc-

tivity in the apple have in most cases given negative results. WHITTEN (1915) found no effect of selecting scions from a high-yielding or low-yielding tree. The scions were grafted in 1895 so that sufficient time has elapsed for conclusive results.

MACOUN's (1916) data are based on three parental trees and seven progeny from each tree. The following table was prepared from MACOUN's data.

TABLE 11 A

PARENT TREE'S PERFORMANCE	YIELD 5 YEARS	PROGENY BY SCIONS	
		Number of trees	Mean annual yield
Heavy yielder.....	103 gallons	7	6.7± .6
Regular yielder.....	88 gallons	7	9.4±1.5
Poor yielder.....	39 gallons	7	6.8± .5

When comparisons are made in view of the probable errors we find that there is no significant difference in the performance of the trees regardless of their origin.

TABLE 11 B

Yields of three parental trees and yields and trunk circumference of their progeny. Data from DAVIS (1921).

TREE	PARENTAL YIELDS	PROGENY YIELDS	DIFFERENCES IN PROGENY YIELD	DIFFERENCE / P. E. Diff.	n
1 Heavy	104.75 gallons	57.5±2.4	1 and 2= 8.7±4.4	2.0	17
2 Medium	78.75 "	48.8±3.7	2 and 3=13.2±4.4	3.0	12
3 Poor	41.0 "	35.6±2.4	3 and 1=21.9±3.4	6.4	8

TREE		PROGENY CIRCUMFERENCE	DIFFERENCES IN PROGENY CIRCUMFERENCE	DIFFERENCE / P.E. Diff.	n
1 Heavy		15.9±.2	1 and 2=1.0±.36	2.8	17
2 Medium		14.9±.3	2 and 3=1.1±.36	3.1	12
3 Poor		13.8±.2	1 and 3=2.1±.28	7.5	8

Recently DAVIS (1921) has presented additional data from MACOUN's scion-selection experiment. Results covering a nine-year period show

some correlation between performance of parents and progeny. As indicated in table 11 B the differences in performance of progeny from the high- and regular-bearing parents may be of questionable significance, but the progeny of the unproductive parent are significantly unproductive compared with the other trees. The fact that in the unproductive progeny the trees are not thrifty, the death rate is high, and the trees are small, may indicate that a disease transmitted in budding is the cause of the unproductive trees. The relatively high yields of the "unproductive" progeny during the early years of the bearing period could then be explained on the basis that stunted or unthrifty trees are often forced into bearing earlier than vigorous trees. It is also significant that of the 17 progeny of the productive parent two trees yielded less than the average yield of the unproductive progeny. But in any event whether the differences in productivity are due to bud mutation or to other factors the results show the advisability of selecting buds or scions from large vigorous apple trees, a suggestion which is perhaps unnecessary, as no one would select scions or buds from weak or unhealthy trees.

Recently CUMMINGS (1921) has reported the results of a rather extensive bud-selection experiment in Vermont. Scions were selected from productive and unproductive trees of seven varieties. Two hundred and forty-eight scions were used, 120 scions from productive trees and 128 from unproductive ones. At the end of ten years 86 of the "productive" and 82 of the "unproductive" scions had borne fruit. In many cases the two classes of scions were grafted into the same variety or even into the same tree. In general the scions from the productive trees were no more productive, in fact they were somewhat less productive, than scions from the unproductive trees. The difference in favor of the "unproductive" scions is probably not significant, but at least there is no indication that scions from productive trees are superior to scions from unproductive trees of a clonal variety. In other words the differences in performance of trees of a clonal variety of apples, in the above case at least, are apparently not inherent, but are the result of environmental factors.

One of the most critical tests of the efficacy of bud selection has been made by RAWES (1922) in England. In 1913 buds were selected from productive and unproductive trees of four varieties of apples. The parental trees were free from disease and were under the same environmental conditions. Buds and scions were worked on crab and dwarfing stocks. The stocks were selected for uniformity. A total of 130 trees were grown, the progeny of productive and unproductive parents being grown in alternate rows. The average number of fruits per tree borne by

the "good history" and "bad history" trees, respectively, for each of the four varieties is as follows: 52 and 54, 78 and 76, 82 and 78, and 74 and 74. In one case individual tree yields are given. The twenty-five "good history" trees bore an average of 52.6 ± 4.3 apples per tree while the "bad history" trees bore 53.8 ± 7.6 apples per tree. The difference in productivity, 1.2 ± 8.8 is not significant. In no case is there any significant difference in the performance of progeny of productive and unproductive trees in any of the four varieties tested. Variability of trees within each class is attributed to the variation in stocks. RAWES emphasizes the necessity of selecting buds and scions from healthy trees.

EAST'S (1910) work indicates that there is no effect of selecting high-yielding hills of potatoes in a clonal variety. An analysis of WHIPPLE'S (1920) data shows that when diseased strains are eliminated there are no permanent differences in the performance of tuber lines of a clonal variety of potatoes.

The recent work of SHAMEL and his colleagues has been generally accepted as proof that differences in productivity within a clonal variety are due to bud mutations and can be transmitted by selected buds. In reporting on the work of SHAMEL and his associates, E. T. MEREDITH (1921, page 392) states that "a tree propagated from a high-yielding tree is itself, in turn, a larger producer, while one propagated from a tree that is practically barren perpetuates the characteristics of the parent tree." The possibility of improving undesirable trees by top-working has been generally accepted by citrus growers. The practical execution of this idea is seen in California where thousands of unproductive or otherwise undesirable trees have been top-worked with selected buds. The CALIFORNIA NURSERYMEN'S BUD SELECTION ASSOCIATION has supplied more than 1,000,000 buds to growers and nurserymen.

Let us examine the evidence on which the above conclusions are based. According to SHAMEL, SCOTT and POMEROY (1918 a, page 145)

> "More than 40,000 undesirable citrus trees in California have been top-worked with buds selected from the most desirable trees in the investigational performance-record plots. The oldest tops grown from these buds are now more than 5 years of age. So far not a single failure has been observed in transmitting the characteristics of the parent trees by means of selected buds."

Variable fruits, attributed to bud variation, are of rather frequent occurrence even in the most desirable strains of orange and lemon varieties. The percentage of variable fruits found in the best strains as recorded by SHAMEL and his co-workers is 0.66 percent for the Washington Navel

strain, 3.4 percent for the Valencia strain, 34.35 percent for the Lisbon strain, and 7.68 percent for the Eureka strain. These strains of oranges and lemons are the ones from which buds are selected for top-working unproductive and undesirable trees. These variable fruits are attributed to bud variation and it would appear that bud variations even in the most desirable strains from which buds are selected, are comparatively frequent. In view of the number of bud variations that are said to occur in the parental trees it is most remarkable that in more than 40,000 top-worked trees, involving over 160,000 buds, not a single bud mutation has been observed.

The conclusions drawn by SHAMEL, SCOTT, POMEROY and DYER (1920 b, page 11), in regard to the occurrence of variable fruits are of interest as they show the inadequacy of the methods used in analyzing the data presented. The statement is made that "usually the most productive trees show the fewest marked variations in fruits from the type of the strain to which they belong." A similar correlation is said to exist in all of the citrus fruits studied. From this statement we might conclude that bud variations are less frequent in high-yielding trees from which buds are selected. However, when we determine the correlation between yield and percentage of variable fruits for the 151 Washington Navel orange trees recorded for 6 years we find that $r = .01 \pm .05$. It is evident that there is absolutely no significant correlation between yield and variability of fruit in the Washington Navel orange trees recorded.

The data for the Lisbon lemons are obviously heterogeneous. The correlation between percentage of variable fruit and yield taken as a whole is practically zero ($r = -.11 \pm .06$). If we omit the 8 "dense unproductive" trees, which are very low yielders and bear a high percentage of variable fruit, we find for the correlation between yield and variable fruit that $r = .68 \pm .04$. The high-yielding trees clearly bear the largest percentage of variable fruits and yet SHAMEL and his colleagues conclude that "usually the most productive trees show the fewest marked variations in fruits."

Top-worked citrus trees begin bearing in two or three years and reach normal production in four or five years. Many of the trees top-worked by SHAMEL have been producing for five or six years, but as yet no data have been presented on the performance of these trees. Until data on the performance of the top-worked trees are presented we are not justified in concluding that differences in productivity are transmitted by selected buds. It is possible that an unproductive tree or one which produces undesirable fruit will become more productive and bear better fruit simply

because it is cut back severely. In order to test the transmission of productiveness through buds or scions it is not enough to top-work low-producing trees with buds or scions from high-producers. High-producing trees should be top-worked with buds from low-producing trees, and *vice versa*. The numbers of trees involved should be large enough to eliminate chance differences in performance, and records of the top-worked trees should be kept. *Under such conditions if buds or scions from low-producing trees remain low-producers when top-worked on high-producing trees, and if buds from high-producers remain high-producers when top-worked on low-producing trees, we would be justified in concluding that differences in productivity are inherent and due to bud variation.* At present only conflicting or unsatisfactory evidence has been presented to support such a conclusion.

CONCLUSIONS

There is great variability in productivity of individuals within clonal varieties of apple and citrus trees. The average coefficient of variation for yield was found to be 70 for Ben Davis apple trees in Maine, 28 for Rome trees in New York, 41 for Jonathan apple trees in Utah, and from 36 to 54 in certain citrus varieties in California. The differences in performance of individual trees were relatively permanent, as indicated by the high value of the interannual correlations which varied from .51 for Rome apple trees to .74 for selected Washington Navel orange trees.

Our problem is to determine what part of the consistent differences in productivity are due to environmental conditions and to what extent these differences are due to inherent differences in individual trees. The important known factors which may cause variability in trees of a clonal variety are: (1) soil heterogeneity, including the effects of soil moisture, elevation, and exposure; (2) differences in vigor and compatibility of seedling stocks on which the trees are grafted; and (3) inherent differences in scions or buds of a clonal variety due to bud mutation.

Soil heterogeneity was found to cause considerable variability in yield of trees. A soil heterogeneity of .40 in our Ben Davis orchard was reduced to $-.07$ by selecting a uniform block of 95 trees and as a result the coefficient of variability was reduced from 58 to 43 in 1914. A reduction in soil heterogeneity also reduced the variability of Navel and Valencia orange trees in California (see table 9). The variability of soil is also shown by a coefficient of variability of 47 for 72 blocks of 16 trees each in our Ben Davis orchard and the relative permanence

of these differences is indicated by an average interannual correlation of .47. Consistent differences in tree performance have apparently been considered by some investigators as proof of inherent differences in trees of a clonal variety, but the fact that permanent differences in yield have been found in *plots* of apple trees, and permanent plots of millet, alfalfa and maize, should discredit any such idea (see table 8).

On the other hand, much variability in tree performance may exist even on soil that is, in general, homogeneous. On soil that was found to be uniform when tested for soil heterogeneity the coefficient of variation was found to be 43 for Ben Davis apple trees, 35 for Jonathan apple trees in Utah, 26 for Rome apple trees in New York, 48 for seedling walnuts, and 28 for Navel oranges in California (table 9). Of even greater importance is the fact that these differences in productivity are just as permanent as the greater differences found on highly heterogeneous soil. For example the interannual correlation was found to be .56 for 882 Ben Davis trees on heterogeneous soil ($r = .40$), while for 95 trees on uniform soil ($r = -.07$) the interannual correlation was .65. In other words soil heterogeneity has no effect on the *permanence* of permanent differences in yield. This may sound like a paradoxical statement, but a little consideration will show its validity. If, for instance, all the trees in our Ben Davis orchard were placed in absolutely uniform soil the permanence of differences in tree productivity would be just as great as before although the variability of total yield would be reduced. It is impossible to perform such an experiment, but it is possible to eliminate the *effect* of soil differences by holding the factor of soil variability constant. This was done, and yet there was no reduction in the *permanence* of differences in yield (see table 10).

Since variability in tree productivity may exist on uniform soil and since soil differences had no effect on the permanence of differences in yield of trees, other factors must be responsible for the permanence of variability in clonal varieties of fruit trees. One of these factors is the variability of seedling stocks on which the clonal varieties are grafted. Seedling stocks are known to be extremely variable and to cause variability in growth of scions or buds. In our seedling orchard, 586 trees which were planted in 1911, vary in circumference from two to eighteen centimeters and the coefficient of variability was found to be 32 percent. In our "stock and scion" orchard we find a high correlation between the size of the tree shortly after it was set in the orchard and its size six years later. All buds of each variety were from a single tree, thus eliminating any possible effect of bud mutation. Thus the consistent differences in

size can often be attributed to variable stocks. WEBBER (1920) has shown the effect of variable stocks on growth of citrus trees in California. Although the work on stocks is limited, there is sufficient evidence to show that their variability is an important factor in causing differences in yield of trees of a clonal variety.

Variability in productivity of trees of a clonal variety have also been attributed to bud variation. The paucity of bud mutations affecting qualitative characters in the apple, and the bud selection work of WHITTEN (1915), CUMMINGS (1921), and RAWES (1922), does not support the belief that bud mutation is an important factor in causing variation in performance of apple trees. The bud-selection work of MACOUN recently reported by DAVIS (1921) does, however, show the transmission of high and low productivity to the progeny by budding. Whether these differences in productivity are due to bud mutation or to the transmission of disease may well be questioned. DAVIS believes that buds from a relatively unproductive but vigorous healthy tree, would give progeny as productive as buds selected from productive vigorous trees. At any rate the evidence that buds should be selected only from vigorous trees does not necessarily mean that bud mutation is a factor in causing variability within a clonal variety.

As the result of the work of SHAMEL and his co-workers bud variations influencing productivity have been generally considered to be of common occurrence in citrus varieties. Although these writers do not directly state in the papers cited, that buds from productive trees result in productive progeny and that buds from unproductive trees produce unproductive progeny, they do state that "all of the strains described can be isolated through bud selection" and in most cases the "strains" are described as "heavy bearers," "less productive," or "unproductive," as compared with the original or standard strain of the variety. SHAMEL and his colleagues may have evidence to prove that in some cases differences in productivity are due to bud mutation, but there are no data in any of their publications to date to warrant such a conclusion.

In conclusion we wish to emphasize the relative part played by environmental fluctuating factors and factors which control the permanence in yield of trees of a clonal variety. The variation remaining after the control of any factor is equal to $\sigma_x\sqrt{1-r^2}$. In other words the percentage of variation in yield remaining is equal to $\sqrt{1-r^2}$. The percentage of variation remaining for the permanence in variability is, for the obviously highly selected data on the Washington Navel orange, $\sqrt{1-.74^2}$ or 67 percent, for the Valencia orange 86 percent, for the Lisbon lemon 78

percent, for the Rome apple trees 86 percent and for the Ben Davis apple trees 85 percent (see table 8). Thus, excluding the Washington Navel orange data as being at least of doubtful value, only from 14 to 22 percent of the variability in yield is due to factors causing permanence in yield. Variable stocks and inherent differences in scions are the important known factors which may cause such permanent differences in yield, and if differences in stocks cause much of this remaining variability, certainly other factors, such as bud mutation, can play little part in causing permanent differences in yield of trees of a clonal variety.

The relative effect of various factors in causing variability in yield of trees has been determined for our Ben Davis orchard. About 62 percent of the variability is due to environmental factors other than soil. Soil heterogeneity causes 18 percent of the variability. Only about 15 percent of the tree variability can be attributed to factors which cause permanence of yield, i.e., grafting stocks, bud mutation, and unknown factors.

LITERATURE CITED

BARRUS, M. F., 1918 Potato-mosaic and certified seed. Potato Mag. **1**: 13-14.

BATCHELOR, L. D., and REED, H. S., 1918 Relation of the variability of yields of fruit trees to the accuracy of field trials. Jour. Agric. Res. **12**: 245-283.

CUMMINGS, M. B., 1921 Apple cion selection from high and low-yielding parent trees. Vermont Agric. Exp. Sta. Bull. 221, pp. 36-38.

DAVIS, M. B., 1921 The possibility of the transmission by asexual propagation of the high-yielding ability of individual apple trees. Scientific Agriculture **2**: 120-124.

EAST, E. M., 1910 The transmission of variations in the potato in asexual reproduction. Rept. Connecticut Agric. Exp. Sta. 1909-1910, pp. 119-160.

GOWEN, M. S., and GOWEN, J. W., 1922 Studies in milk secretion XIII. Relation between milk yields and butter-fat percentages of the 7-day and 365-day tests of Holstein-Friesian Advanced Registry Cattle. Maine Agric. Exp. Sta. Bull. 306, pp. 23-60.

GREEN, W. J., and WAID, C., 1906 Potato investigations. Ohio Agric. Exp. Sta. Bull. 174, pp. 251-289.

HARRIS, J. A., 1911 On the formation of correlation and contingency tables when the number of combinations is large. Amer. Nat. **45**: 566-571.

1913 On the calculation of intra-class and inter-class coefficients of correlation from class moments when the number of possible combinations is large. Biometrika **9**: 446-472.

1915 On a criterion of substratum homogeneity (or heterogeneity) in field experiments. Amer. Nat. **49**: 430-454.

1920 Practical universality of field heterogeneity as a factor influencing plot yields. Jour. Agric. Res. **19**: 279-314.

HARRIS, J. A., and SCOFIELD, C. S., 1920 Permanence of differences in the plots of an experimental field. Jour. Agric. Res. **20**: 335-356.

HEDRICK, U. P., 1915 Dwarf apples. New York Agric. Exp. Sta. Bull. 406, pp. 341-368.

HEDRICK, U. P., and ANTHONY, R. D., 1919 Twenty years of fertilizers in an apple orchard. New York Agric. Exp. Sta. Bull. 460, pp. 71-96.

MACOUN, W. T., 1916 The apple in Canada; its cultivation and improvement. Dominion Exp. Farms, Bull. 86, pp. 27-31.

MELHUS, I. E., 1917 Notes on mosaic symptoms of Irish potatoes (Abstract). Phytopath. 7: 71.

MEREDITH, E. T., 1921 The Department's service to fruit growers. Amer. Fruit Grower 41: 6.

PEARL, R., and MINER, J. R., 1919 Variation of Ayrshire cows in the quantity and fat content of their milk. Jour. Agric. Res. 17: 285-322.

RAWES, A. N., 1922 Comparative cropping of apple trees propagated from parent trees of good history and bad history. Jour. Roy. Hort. Soc. 47: 163-168.

SHAMEL, A. D., SCOTT, L. B., and POMEROY, C. S., 1918a Citrus-fruit improvement: a study of bud variation in the Washington Navel orange. United States Dept. Agric. Bull. 623, pp. 1-146.

1918b Citrus-fruit improvement: a study of bud variation in the Valencia orange. United States Dept. Agric. Bull. 624, pp. 1-120.

SHAMEL, A. D., SCOTT, L. B., POMEROY, C. S., and DYER, C. L., 1920 Citrus-fruit improvement: a study of bud variation in the Lisbon lemon. United States Dept. Agric. Bull. 815, pp. 1-70.

1920 Citrus-fruit improvement: a study of bud variation in the Eureka lemon. United States Dept. Agric. Bull. 813, pp. 1-88.

SCHULTZ, E. S., FOLSOM, D., HILDEBRANDT, F. M., and HAWKINS, L. A., 1919 Investigations on the mosaic disease of the Irish potato. Jour. Agric. Res. 17: 247-273.

SCHULTZ, E. S., and FOLSOM, D., 1920 Transmission of the mosaic disease of Irish potatoes. Jour. Agric. Res. 19: 315-337.

STEWART, J. P., 1911 Factors influencing yield, color, size and growth in apples. Pennsylvania State Coll. Rept. 1910-11, pp. 401-511.

STUART, W., 1915 Potato breeding and selection. United States Dept. Agric. Bull. 195, pp. 1-35.

SURFACE, F. M., and PEARL, R., 1916 A method of correcting for soil heterogeneity in variety tests. Jour. Agric. Res. 5: 1039-1050.

WARING, J. H., 1920 The probable value of trunk circumference as an adjunct to fruit yield in interpreting apple orchard experiments. Amer. Soc. Hort. Sci. 1920: 179-185.

WEBBER, H. J., 1920 The improvement of root-stocks. Jour. Heredity 11: 291-299.

WHIPPLE, O. B., 1920 Line-selection work with potatoes. Jour. Agric. Res. 19: 543-573.

WHITTEN, J. C., 1915 Bud selection for increasing yields. Missouri Agric. Exp. Sta. Bull. 131, pp. 479-480.

INTERSPECIFIC HYBRIDS IN CREPIS. II. A PRELIMINARY REPORT ON THE RESULTS OF HYBRIDIZING *CREPIS SETOSA* HALL. WITH *C. CAPILLARIS* (L.) WALLR. AND WITH *C. BIENNIS* L.

J. L. COLLINS AND MARGARET C. MANN
University of California, Berkeley, California

Received December 14, 1922

TABLE OF CONTENTS

	PAGE
INTRODUCTION	212
Material and methods	213
Crepis setosa Hall. × *C. capillaris* (L.) Wallr.	214
The F_1 hybrid	214
Back-crosses	217
Crepis setosa Hall. × *C. biennis* L.	221
The F_1 hybrids	221
The F_2 plants	227
Back-crosses	227
SUMMARY AND CONCLUSIONS	228
LITERATURE CITED	232

INTRODUCTION

In the first paper of this series (BABCOCK and COLLINS 1920) the F_1 hybrid of *Crepis capillaris* (L.) Wallr. (N = 3)[1] × *Crepis tectorum* L. (N = 4)[1] was described in detail. It was pointed out that the compatibility between these two species was so low that the two haploid sets of chromosomes were unable to function together, with the result that the F_1 plants were unable to grow beyond the seedling stage. The present paper presents data on two other species hybrids obtained by crossing *Crepis setosa* Hall. with *C. capillaris* (L.) Wallr. and with *C. biennis* L., both of which are able to complete growth to maturity. These hybrids are of especial interest because they involve (1) a cross between two species which differ from one another by one pair of chromosomes, *setosa* (N = 4) × *capillaris* (N = 3), and (2) a cross between two species differing by 16 pairs of chromosomes, *setosa* (N = 4) × *biennis* (N = 20). Although back-cross and F_2 plants have been obtained, this report is necessarily preliminary, since anything like a complete analysis of the problems which such material involves will occupy a considerable period.

[1] N is the haploid number of chromosomes.

MATERIAL AND METHODS

The oriental species of the genus have been divided by BOISSIER (1875) into three sections largely on the basis of the achene characters. Section I, Eucrepis, includes those plants having achenes all similar, the apex sometimes more or less narrowed or attenuated but never beaked. Both *Crepis capillaris* and *C. biennis* are in this section. Section II, Barkhausia, includes the plants having all or at least the central achenes beaked, the pappus being borne at the end of the slender beak. *Crepis setosa* belongs in this section. Section III, Gatyona, includes species having achenes either beaked or non-beaked, but placed radially with a conspicuous wing or keel on the ventral face of the achene. None of the species reported upon are representatives of this third section.

Crepis capillaris (L.) Wallr. is an annual, producing a stem, single below and paniculately branched above or branching from the crown of the root as well as along the stem. The plant varies in height from a few inches to four feet. The leaves are generally considered as glabrous but in reality have a fine scattered pubescence over the surface which becomes much more pronounced on the midrib of the radical leaves. These leaves are generally lanceolate in form, the margins of those on different plants varying from entire or merely toothed to deeply pinnatifid or bipinnately compound. The cauline leaves are sagittate or sub-sagittate. The involucre is composed of two series of foliaceous bracts the outer of which are smaller, forming an appressed calyculus. The bracts of the inner series are larger and enclose the florets. The keel of these bracts is covered with short gland-tipped hairs.[2] The achenes are 10-ribbed, tapered at both ends, without a beak, length about 2.5 mm (figure 1 b). The pappus is pure white.

Crepis biennis L. is also in the Eucrepis section. It is a biennial under ordinary conditions. The plant is larger and coarser in all parts than the corresponding parts in *C. capillaris* but differs very little in morphological characters, when size is not considered. Erect radical leaves are characteristic of our *C. biennis* material, while they are seldom seen in *C. capillaris*; the cauline leaves in *C. biennis* are not sagittate. On the other hand, the stem habit is approximately the same for both species, but the plant-pubescence and leaf-form variation exhibit some differences, which are in keeping with the larger size. The achenes of *C. biennis* are approximately 3.5 mm in length. The principal features of difference

[2] Plants lacking this glandular pubescence are known to act as Mendelian recessives when crossed with the normal type.

between the two species, then, is found in the size and in the length of the life cycle of an individual.

Crepis setosa Hall. is in the Barkhausia section, since the achenes are beaked. It is an annual of about the same height and growth as *capillaris*. Stems and leaves are bristly pubescent but not glandular, the bristle-like hairs being conspicuous on the pedicel and involucre. The calyculus is relaxed or spreading (figure 1 a). Radical leaves show considerable variation in degree of pinnate lobing. Cauline leaves are sagittate at the base. The achenes are more slender than in *C. capillaris*; they are usually 10-ribbed and are striate-tuberculate, with the pappus at the tip of the slender beak (figure 1b).

Two methods were used to secure cross-fertilization. One consisted in emasculation of ten to fifteen florets per head, the remaining floret buds being removed. The other method consisted in washing the fresh pollen from the stigmas of the peripheral whorl of florets with a fine stream of water from a syringe; at the same time the central unopened florets were removed. Pollen was applied from one to three days later. In both F_1 hybrids described *C. setosa* was used as the female parent.

The achenes were placed in moist-chamber germinators. When germination had occurred the seedlings were transferred to six-inch clay pots where they remained in the greenhouse. (Cf. COLLINS 1922 for detailed description of the methods employed.)

For chromosome study root tips were fixed in chrom-acetic-urea and stained in Heidenhain's iron haematoxylin. The root tips were usually obtained from the bottom of the pots, but some were from sprouted achenes. BELLING's (1921) iron-aceto-carmine method proved very useful for pollen mother cells of Crepis.

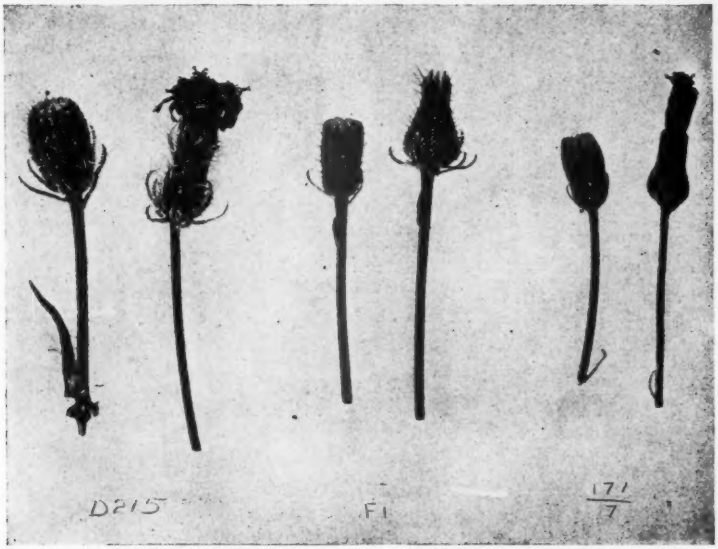
Crepis setosa HALL. × *C. capillaris* (L.) WALLR.

The F_1 hybrid

The *setosa* characters predominate in F_1 with the one marked exception that the achenes are not beaked (figure 1 b). The rosette leaves of the hybrid have the pinnate lobing intensified (figure 4) while the pubescence throughout the plant is less pronounced than in *C. setosa*.

Crepis setosa, which was used as the pistillate parent in the cross, has four pairs of chromosomes with such marked individuality that there is almost never any difficulty in identifying the homologues in somatic cells. The longest chromosomes always have a semi-detached tip which frequently makes them look like i's or j's. They are readily found in the

figures (figure 2 c). The other pairs consist of a very short pair and two pairs of intermediate size, one slightly longer than the other.



a

b

FIGURE 1.—a. Flower heads of *Crepis setosa* (left), *C. capillaris* (right) and F_1 hybrid (center). b. Achenes of *Crepis setosa* (left), *C. capillaris* (right) and F_1 hybrid (center).

Crepis capillaris has three pairs of chromosomes which also show distinct individuality (figure 2 b) as was previously reported by DIGBY (1914),

ROSENBERG (1909) and SAX (1918).[3] Here the chromosomes consist of a graded series, one long, one intermediate, and one short pair. The longest

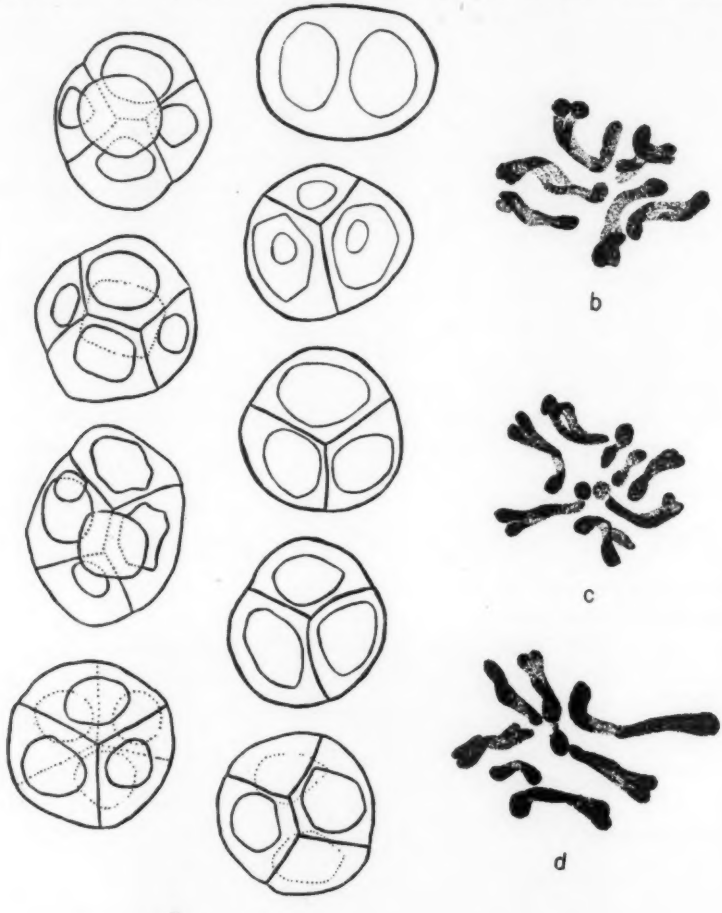


FIGURE 2.—a. Anomalies in pollen formation in the F_1 *Crepis setosa* × *C. capillaris*. b. Somatic metaphase of *C. capillaris*. c. Somatic metaphase of *C. setosa*. d. Somatic metaphase of *C. setosa* × *C. capillaris*.

chromosome of *C. capillaris* is decidedly longer than the longest of *C. setosa*, and the shortest of *C. capillaris* is much longer than the shortest

[3] Unpublished.

of *C. setosa*. There is little difference in the size of the two intermediate pairs of *C. setosa* and the two shorter pairs of *C. capillaris*.

The F_1 root-tip cells have seven chromosomes, four from *C. setosa* and three from *C. capillaris* (figure 2 d). It is nearly always possible to recognize the long compound one with the partially detached end and the very short chromosome from the *setosa* complex, as well as the longest one from *C. capillaris* but it is impossible to distinguish the two shorter *capillaris* from the two intermediate *setosa* chromosomes. The chromosomes in the pollen mother cells are unpaired in diakinesis and seem to undergo rather less contraction than is usual at this stage (figure 4 a). The first meiotic division is very irregular. The undivided chromosomes are usually distributed in a variable number of large or small groups which are scattered throughout the pollen mother cell and there form from 2 to 6 vesicles or nuclei, a condition which is frequently found in cells under abnormal conditions. In a single case all of the chromosomes had divided, but this is the exception, not the rule. The second division looks much more normal than the first, but it appears to occur only when the chromosomes have previously divided into two rather equal groups. The cytoplasmic divisions usually give rise to the normal tetrad, although about 12 percent result in two, 19 percent in three, 1.4 percent in five and 0.7 percent in six microspores. When three, four, five or six microspores result, they may be very similar or very different in size, as may be seen from figure 2 a. The size of the cell corresponds to the size of the nucleus which it contains.

Back-crosses

Very little of the pollen appears to mature properly, most of it remaining in the anther locules and there disintegrating. It is to be expected that some of the pollen would be functional yet none of the back-crosses in which the F_1 was used as the pollen parent was successful. That some ovules are functional has been shown by the fact that back-crosses were obtained when the F_1 was used as the pistillate parent. Five such plants are now growing, two of them having 10, two 8, and one 7 chromosomes in the root-tip cells. The different types of gametes which are to be expected from a random assortment of the seven chromosomes of the F_1 hybrid during germ-cell formation are shown in table 1. The relative number of each expected type is shown in the first column. All of the 128 gametes are different even though many of them have the same chromosome number. It will be noted in column five that three of the five back-cross plants which were obtained, the two with 8 and the one with 7

chromosomes, belong to the two classes containing the largest number of different gametes. The other two plants must have resulted from the fertilization of egg cells with six chromosomes.

Only one of the back-crosses has flowered, $22.306P_1$ (figure 3 b). Four pairs of chromosomes appear in diakinesis, and reduction, cytokinesis, and pollen formation proceed normally. This is in accord with the observation that the root-tip cells have the same type of chromosome group as *C. setosa* and that the plant is typically *C. setosa* in appearance. It may thus be expected to maintain itself as such upon inbreeding. The other back-cross with eight chromosomes, $22.307P_3$ (figure 3 a), has a very similar type of chromosome group in the somatic cells, and the rosette leaves are well within the range of variability which has been observed

TABLE 1

Showing the chromosome relations when the F_1 setosa × capillaris is back-crossed to setosa.

RELATIVE FREQUENCY OF F_1 GAMETES	NUMBER OF CHROMOSOMES IN			NUMBER OF ZYGOTES OBTAINED
	F_1 gametes	*setosa* gametes	resulting zygote	
1	7	4	11	
7	6	4	10	2
21	5	4	9	
35	4	4	8	2
35	3	4	7	1
21	2	4	6	
7	1	4	5	
1	0	4	4	
128				

in *C. setosa*. Until it has flowered, however, we cannot be certain that all of the chromosomes of intermediate length belong to the *setosa* complex. Perhaps the slowness to flower indicates the presence of one or two *capillaris* chromosomes. The plant with seven chromosomes ($22.307P_1$, figure 4) resembles the F_1 both in rosette-leaf characteristics and in type of somatic chromosome group (cf. figures 4 b and 2 d). The fact that the entire *setosa* group reappeared together at least once and probably twice and that the F_1 type seems to have reappeared among the back-crosses, indicates that back-crosses to *C. capillaris* should give some pure *C. capillaris* plants. Thus far all attempts to make this back-cross have been unsuccessful, but all of the work was done in the greenhouse during the summer, where achenes do not set as readily as they do out-of-doors.

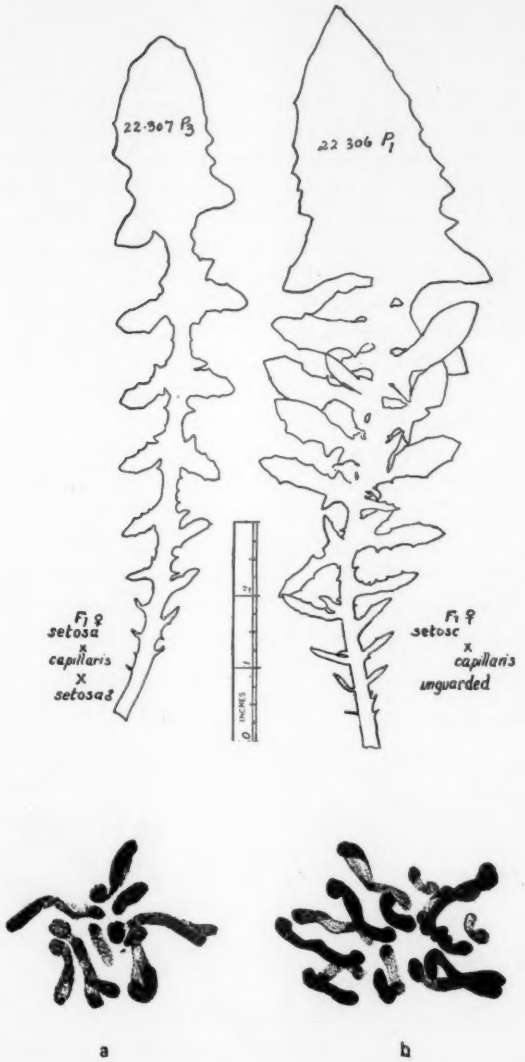


FIGURE 3.—Back-crosses with eight somatic chromosomes. a. Metaphase of 22.307P_3. b. Early anaphase of 22.306P_1. Note correspondence of size and position of divided chromosomes.

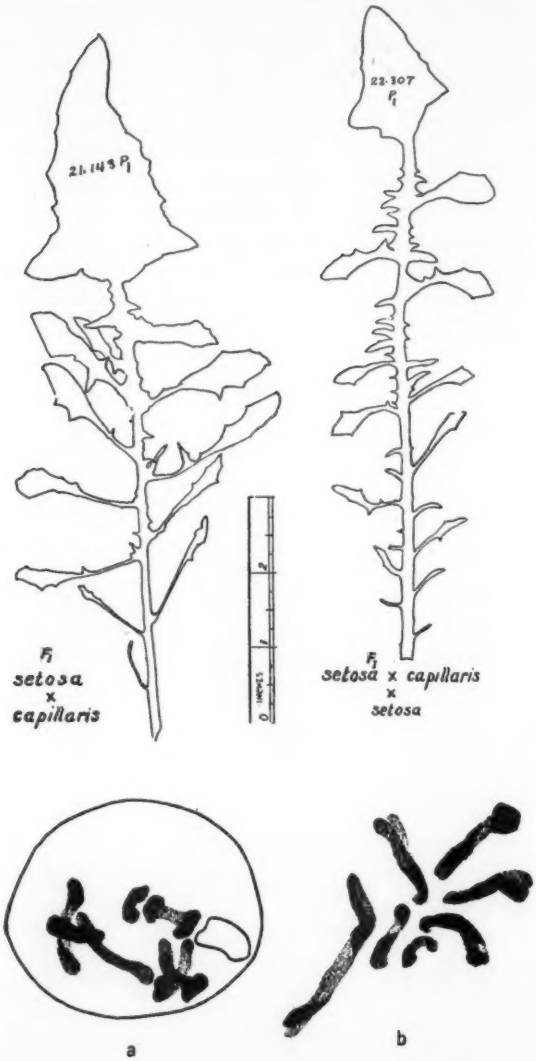


FIGURE 4.—a. Diakinesis group of a pollen mother cell of the F_1 *C. setosa* × *C. capillaris*. The nucleolus is outlined, the chromosomes stippled. b. Somatic group of (*C. setosa* × *C. capillaris*) × *C. setosa*, 22.307P_1.

The two plants containing 10 chromosomes in the root-tip cells (figure 5) were strikingly different in the rosette stage from any of the other plants, but very similar to one another. They are only in the rosette stage at present and have not yet flowered, hence nothing is known regarding their meiotic behavior and fertility. The somatic groups contain two j's and two short *setosa* chromosomes, as well as one long one from *C. capillaris*. The source of the other five chromosomes is still doubtful. The meiotic behavior in the plants resulting from back-crossing the F_1 to *C. setosa* should be of great interest, since one would expect to be able to determine positively the number of unpaired chromosomes in diakinesis. The number of lagging chromosomes in the first meiotic division would also assist in the determination, since univalent chromosomes usually lag while bivalents divide at a more rapid rate. If one of the 10-chromosome back-cross plants were to show only 2 unpaired chromosomes it would indicate that the F_1 gamete contained 4 *setosa*, a haploid set, and 2 *capillaris* chromosomes. If, on the other hand, four unpaired chromosomes were found it would indicate that the F_1 gamete was composed of 3 *capillaris*, a complete haploid set, and 3 *setosa* chromosomes. If the 10-chromosome plants contain 8 *setosa* and 2 *capillaris* chromosomes the gametic series with respect to number and derivation of chromosomes (S = *setosa* and C = *capillaris*) will be 1 (4 S+0 C)+2 (4 S+1 C)+1(4 S+2 C) and the zygotic series produced by self-fertilization of such a plant will be:

1 (8 S+0 C)
4 (8 S+1 C)
6 (8 S+2 C)
4 (8 S+3 C)
1 (8 S+4 C)

Those plants which have one or two pairs of *capillaris* chromosomes plus the four pairs of *setosa* chromosomes would be expected to continue as 5- and 6-pair species, respectively, in subsequent generations.

Crepis setosa HALL. × *C. biennis* L.

The F_1 hybrids

The plants of the F_1 are predominately like the pollen parent, *C. biennis*, but differ in having finer pinnate lobing of the rosette leaves than is the case in *C. biennis*, and in blooming the first year. The hybrids are quite as large and vigorous as the parent which they resemble. Characteristics of *C. biennis* have been given above.

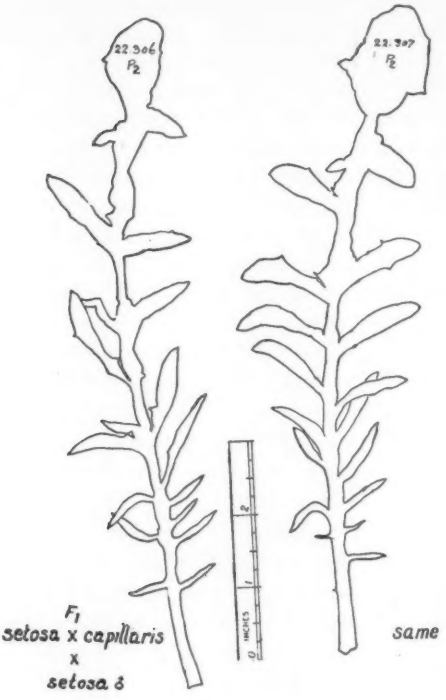


FIGURE 5.—Back-crosses with ten somatic chromosomes. a. Somatic group of 22.307P_2. b. Same of 22. 306P_2.

C. biennis is an outstanding species in the genus in that it has twenty pairs of chromosomes. They are similar in breadth and in length to the chromosomes of most of the other species so that the great increase in number must have resulted from duplication and not from cross-division of the chromosomes. Except for its greater size *C. biennis* does not differ from the four- and some of the five-chromosome species, so that one is tempted to consider it, at least tentatively, as an octoploid species which may have originated in much the same manner as the *gigas* forms of Oenothera and Primula. Such forms have been observed to arise suddenly, as tetraploids, from the normal diploid species, and, at least in certain cases, to maintain the increased chromosome number as well as the *gigas* character in subsequent generations. ROSENBERG (1920), who reported forty-two chromosomes for this species, considered three pairs as a basic number and hence *C. biennis* as a case of 14-ploidy. MARCHAL (1920) counted only sixteen pairs in the pollen mother cells. Counts of the pollen mother cells and root tips of the pure *C. biennis* of our material show that twenty pairs are present in all diploid cells (figure 6). It has been possible to check this count by still another means since a hybrid was obtained between *C. setosa* (N = 4) and *C. biennis* (N = 20), in which the root-tip cells contain twenty-four chromosomes, and since four came from *C. setosa*, twenty must have been derived from the *biennis* parent (figure 7 b 2).

The large number of chromosomes in *C. biennis* makes a study of the individuality impracticable at the present time, but it would be of great interest to know whether or not they fall into five size-types as should be the case if it were an octoploid form derived from a five-pair species.

No other hybrid has been described in which there is anything like as great a difference in the chromosome numbers of the parents as there is in *C. setosa* and *C. biennis*. It has frequently been assumed that such crosses would prove impracticable. The fact that these species will hybridize is astonishing enough, but it is especially surprising to find that the divisions of the pollen mother cells are very nearly normal, much more so than in the cross, *C. setosa* (N = 4) × *C. capillaris* (N = 3). The microspore groups are invariably tetrads and one rarely finds evidence that chromosomes have been cast out into the cytoplasm. However, a few small extra-nuclear vesicles were noted in pollen mother cells which had not yet undergone cytokinesis.

In most hybrids between species which differ in chromosome number, it has been found that the maternal chromosomes tend to pair with the paternal, leaving the extra chromosomes as univalents. TÄCKHOLM (1922) after

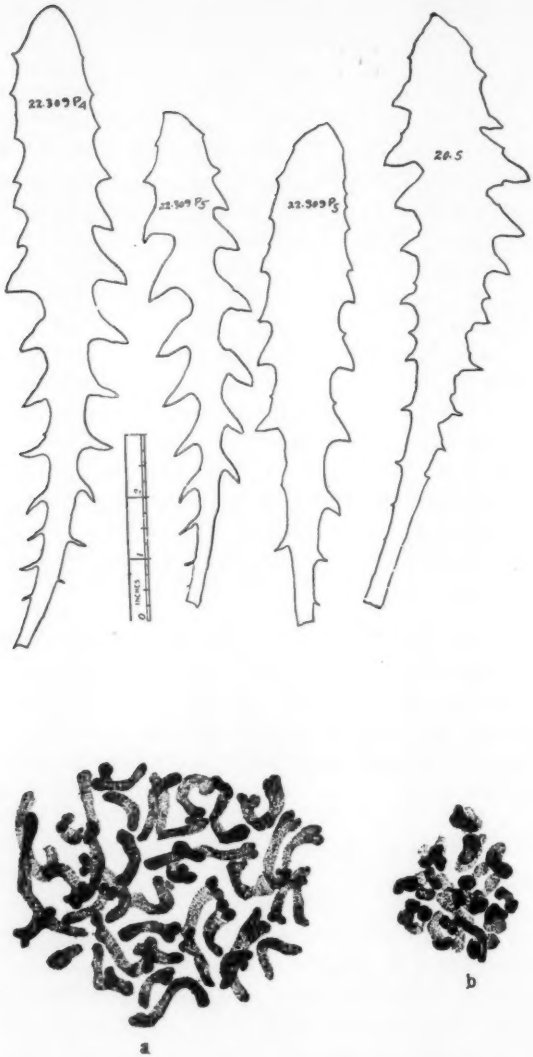


FIGURE 6.—*Crepis biennis*. Variation in leaf types. a. Somatic chromosome group. b. Haploid chromosome group.

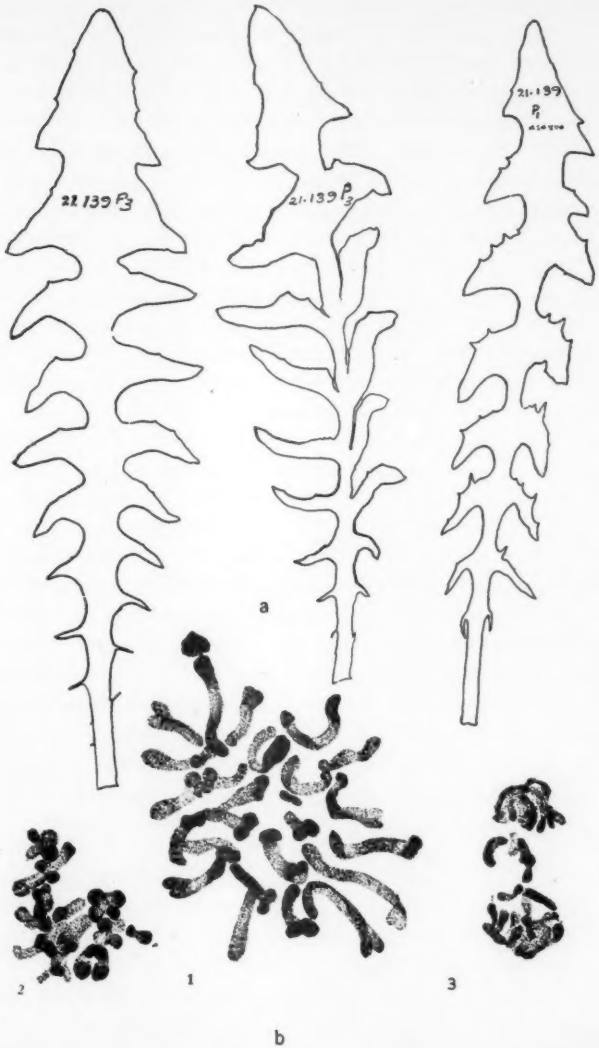


FIGURE 7.—a. Leaf variations in F_1 *C. setosa* × *C. biennis*. b. Chromosome groups of F_1 *C. setosa* × *C. biennis*. 1. Somatic group. 2. Diakinesis showing eight pairs and four odd chromosomes. 3. Same showing four lagging chromosomes which have divided.

examining the literature concludes that this is the rule. He notes that little or no pairing may occur, but that when it does he believes that it has been shown that maternal and paternal chromosomes mate. Only one exception to this appears in the literature and this TÄCKHOLM seriously doubts, namely that of *Digitalis lutea* (N = 48) × *D. pupurea* (N = 24) described by HAASE-BESSEL (1921), which shows 36 pairs in diakinesis. It is obvious that in this case certain of the *D. lutea* chromosomes must have paired with one another. It is of especial interest to note that as many as ten bivalent chromosomes can readily be distinguished in the diakinesis of the pollen mother cells of the F_1 from *C. setosa* × *biennis*. There are also a few obviously univalent chromosomes (figure 7 b). Since *C. setosa* contributes only four chromosomes it is evident that those from *C. biennis* must have paired with one another. This is important, for it indicates a qualitative similarity between what one would usually consider non-homologous chromosomes in *C. biennis*; exactly what one would expect to find if the increased number had been derived by chromosome duplication from a progenitor with a lesser number. The bivalents of the F_1 pollen mother cells are very similar to the bivalents seen in *C. biennis*. At the anaphase of the first division of the pollen mother cells two to four chromosomes of variable size are seen to lag (figure 7 b 3), but eventually, except in rare cases, all of the chromosomes enter one or the other of the two nuclei. The second division is more regular. It seems significant that the number of laggards should not exceed four and that they should show distinct size differences, since this is just what one would expect if the *setosa* chromosomes failed to pair with one another. And since they did not pair in the cross, *C. setosa* (N = 4) × *capillaris* (N = 3) there is reason to suppose that they are qualitatively so different that it is impossible for them to pair together. The lagging chromosomes in the *setosa* × *biennis* hybrid sometimes appear to divide, but sometimes seem also to pass undivided to one or the other pole. It is therefore obvious that one should expect gametes to be formed containing 10 *biennis* chromosomes plus 0, 1, 2, 3 or 4 *setosa* chromosomes, provided there is random assortment of the unpaired *setosa* chromosomes. We have shown below in outline the types of progeny expected in the F_2 generation. It will be seen that five types would have no unpaired chromosomes; these should be expected to continue as 10-, 11-, 12-, 13- and 14-paired races. If these could again be crossed with *C. setosa* other interesting types should be produced containing 5, 6, 7, 8 and 9 pairs of chromosomes, if our assumption that *C. biennis* is an octoploid species is correct. If there is independent assortment of the univalent *setosa*

chromosomes, as now appears to be the case, the gametes of the F_1 hybrid will conform to the following series:

1 (10 B+0 S)
4 (10 B+1 S)
6 (10 B+2 S)
4 (10 B+3 S)
1 (10 B+4 S)

If these gametes all survive and unite at random in fertilization, the F_2 zygotic series will be:

1 (20 B+0 S)
8 (20 B+1 S)
28 (20 B+2 S)
56 (20 B+3 S)
70 (20 B+4 S)
56 (20 B+5 S)
28 (20 B+6 S)
8 (20 B+7 S)
1 (20 B+8 S)

The F_2 plants

The F_2 plants which have been examined cytologically have 24 or 25 chromosomes (figure 9 b). It is significant that these belong to the largest of the theoretical classes outlined above, although whether this means anything in terms of gamete survival is not clear. The five *setosa* chromosomes must include one or two pairs, the chances being about equal for both alternatives. If these mate it may be possible (1) to reduce the race to a 10-pair species, or (2) to continue it as an 11- or 12-pair species (i.e., 10 pairs of *C. biennis* plus 1 or 2 pairs of *C. setosa*).

Back-crosses

The back-cross of the F_1 to *biennis* presents no difficulty. We should expect different types of plants with respect to somatic chromosome number, since the distribution of the *setosa* chromosomes is somewhat irregular in the F_1. They should contain 30, 31, 32, 33 or 34 chromosomes. The two back-crosses which have been studied cytologically have 32 chromosomes in root-tip cells (figure 8 e), twenty from *biennis* plus 10 *biennis* and 2 *setosa* chromosomes from the F_1. It is conceivable that 15 pairs plus 2 laggards, or 10 pairs and 12 laggards might be found in reducing cells, the behavior depending upon whether or not the 10 unmated *biennis* chromosomes can form 5 pairs. Study of pollen mother cells shows that

fifteen pairs are present and chromosome division and cytokinesis proceed normally. Thus we shall expect a 30-chromosome form to result from back-crosses.

It has been more difficult to secure viable achenes when the F_1 is back-crossed to the *setosa* parent. Three such plants have been obtained, all having a different appearance in the rosette stage. One, $22.315P_2$, would pass readily for *biennis*, while $22.315P_1$ closely resembles the F_1 *C. setosa* $\times$ *C. biennis* and $22.315P_3$ is unlike anything which we have observed in the parent species. There are 17 chromosomes in the root-tip cells of $22.315P_2$ (figure 8 d), presumably 10 *biennis* and 3 *setosa* chromosomes having come from the F_1 while a normal haploid set of 4 has entered from *C. setosa*. Inasmuch as 10 *biennis* chromosomes give the *biennis* type of plant it seems probable that they contain all of the genes necessary for its development. There should thus be 3 pairs of *setosa* chromosomes. If the 10 *biennis* chromosomes form 5 pairs and the 6 *setosa* chromosomes 3 pairs, we should be able to obtain an 8-paired species plus one laggard. The other back-crosses, $23.315P_1$ and $22.315P_3$, have not yet been studied cytologically. Back-crosses to *C. setosa* should include plants with 14, 15, 16, 17 and 18 chromosomes in somatic cells and by inbreeding might, as indicated above, produce stable races with 10, 12, 14, 16 and 18 chromosomes, respectively.

The fact that the hybrids of these species can be propagated asexually makes it possible to retain any plant, having a particular chromosome complex, for repeated tests and analysis extending over more than one season.

SUMMARY AND CONCLUSIONS

Such material makes it possible to study the effects in the F_1 of (a) extra chromosomes, and (b) decreased chromosome number, on development and somatic appearance.

When the chromosomes are irregularly distributed, as in the meiosis of the F_1 from *C. setosa* $\times$ *C. capillaris*, a means is afforded whereby chromosome number may, by subsequent breeding, be increased by one or a few pairs. Thus, breaks in chromosome number series, such as we find in the genus Crepis, may be caused by hybridization, not necessarily by non-disjunction.

If *C. setosa* $\times$ *C. capillaris* can be inbred it should be possible to observe the effect of one, two, or three pairs of *capillaris* chromosomes on the characters of the plant.

It has often been assumed, that, if a single pair of chromosomes foreign

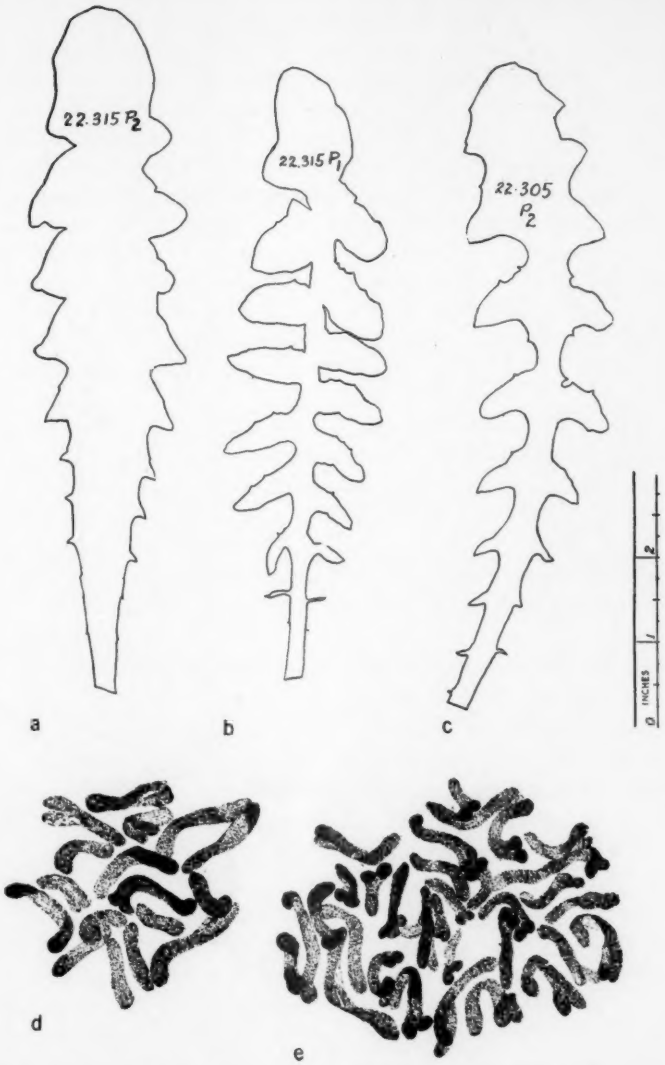
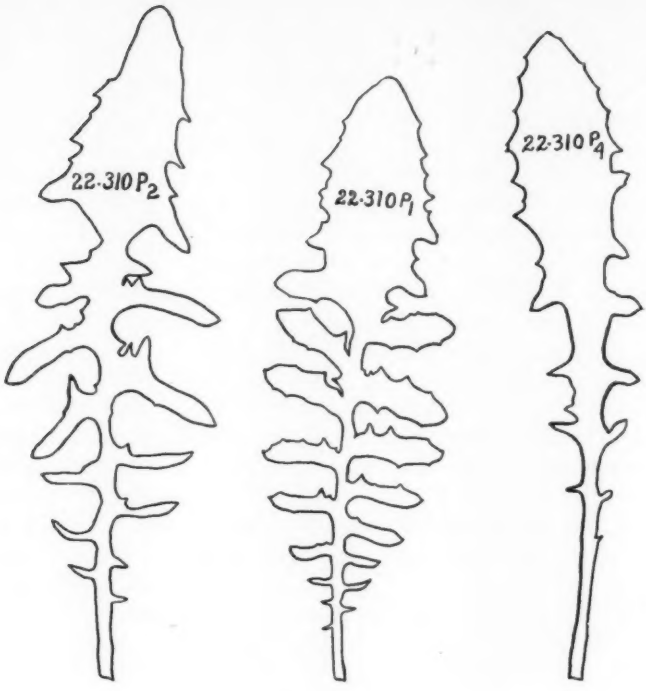
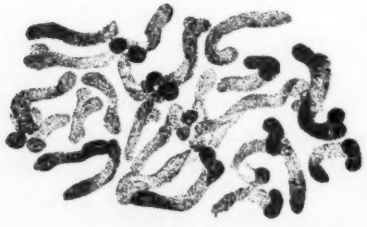


FIGURE 8.—Variations in back-crosses. a and b. F_1 (*C. setosa* × *biennis*) × *C. setosa*. c. F_1 (*C. setosa* × *C. biennis*) × *C. biennis*. d. Chromosome group of 22.315P_2. e. Chromosome group of 22. 304P_2.



a



b

FIGURE 9.—a. Leaf variations in F_2 *C. setosa* × *C. biennis*. b. Chromosome group of F_1 22. 310P_2.

to a group were to be inserted into it, they would automatically maintain themselves. This can now be tested experimentally.

The back-crosses from the F_1 (*C. setosa* × *C. capillaris*) × *C. setosa* indicate that the germ cells of the F_1 have a greater probability of survival if they contain the total *capillaris* or *setosa* complex with or without extra chromosomes.

Since *Crepis biennis* (N = 20) will cross with *Crepis setosa* (N = 4) and give vigorous offspring the possibility of hybridization is obviously not proportionate to the similarity of chromosome number of the parents.

The evidence from *C. setosa* × *biennis* proves that it is possible for the chromosomes from one parent to unite with one another, and shows that plants with half of the normal *biennis* number retain the *biennis* characteristics. This opens the question as to how far this reduction can be carried, and of how such further reduction may affect the somatic appearance.

Since the F_1 of *C. setosa* (N = 4) × *C. capillaris* (N = 3) shows very abnormal reduction phenomena while *C. setosa* × *C. biennis* reduces almost normally, it is evident that normality of reduction does not depend upon similarity of chromosome number, but rather upon likeness of internal composition of the chromosomes.

Plants of *C. biennis* are among the largest and most robust of the genus. The cells are much larger than they are in the species with lower chromosome number.

If *biennis* is considered as an octoploid species the loss of the *gigas* characteristics would not be expected until the 2N condition had been reached. The F_1 *C. setosa* × *C. biennis*, having the *biennis* characters but only a haploid number of *biennis* chromosomes, is as large as *C. biennis*. This may indicate that these plants are still in the tetraploid condition.

The F_1 hybrid, having half of the *biennis* chromosomes, is an annual. This indicates either the activity of the *setosa* haploid set of chromosomes or that reduction of the number of chromosomes in this case results in a change from biennial to annual.

The F_1 *C. setosa* × *C. biennis* cannot revert to the parent species through back-crossing because of the pairing of the *biennis* chromosomes of the F_1 during meiosis. The back-crosses to *C. biennis*, which have ten less *biennis* chromosomes than *C. biennis*, show no *setosa* characteristics in the rosette stage.

In these crosses the species contributing the greater number of chromosomes was dominant. In *C. setosa* (N = 4) × *C. capillaris* (N = 3), *C. setosa*

was dominant while in *C. setosa* (N = 4) × *C. biennis* (N = 20) *C. biennis* was dominant.

It is possible to propagate the hybrids and their parents by asexual means.

Self-fertilized seeds were found to set on *C. setosa* × *C. biennis* F_1 hybrids in the garden, but failed to set in the greenhouse. The F_1 *C. setosa* × *C. capillaris* failed to set selfed seed in the greenhouse but was not tested in the field.

LITERATURE CITED

BABCOCK, E. B., 1920 Crepis—a promising genus for genetic investigations. Amer. Nat. **54**: 271-276.

BABCOCK, E. B., and COLLINS, J. L., 1920 Interspecific hybrids in Crepis. I. *Crepis capillaris* (L.) Wallr. × *C. tectorum* L. Univ. California Publ. Agric. Sci. **2**: 191-204.

BELLING, J., 1921 On counting chromosomes in pollen mother cells. Amer. Nat. **55**: 573-574.

BOISSIER, E., 1875 Flora orientalis Vol. 3. Lugduni (Lyons).

COLLINS, J. L., 1922 Culture of Crepis for genetic investigations. Jour. Heredity **13**: 329-335.

DIGBY, L., 1914 Critical study of the cytology of *Crepis virens*. Archiv f. Zellforsch. **12**: 97-146.

HAASE-BESSEL, G., 1921 Digitalis-Studien. II. Zeitschr. indukt. Abstamm. u. Vererb. **27**: 1-26.

MARCHAL, E., 1920 Recherches sur les variations numeriques des chromosomes dan la série végétale. Memoirs de l'Acad. Roy. Belgique (classe de sciences) 2nd ser. **4**: 3-108.

ROSENBERG, O., 1909 Zur Kenntniss von den Tetradenteilungen der Compositen. Svensk. Bot. Tidskr. **3**: 64-77.

1920 Weitere Untersuchungen über die Chromosomenverhältnisse in Crepis. Särtryck ur Svensk. Bot. Tidskr. **14**: 319-326.

TÄCKHOLM, G., 1922 Zytologische Studien über die Gattung Rosa. Acta Horti Bergiani **7**: 97-381.

ON THE DISTRIBUTION OF MUTANT CHARACTERS AMONG THE CHROMOSOMES OF *OENOTHERA LAMARCKIANA*

HUGO DE VRIES
Lunteren, Holland

AND

K. BOEDIJN
University of Amsterdam, Amsterdam, Holland

Received December 21, 1922

During the last ten years the number of distinct mutant types of *Oenothera Lamarckiana* has increased considerably. One of us has determined the chromosome numbers for most of them, including those older ones, for which these figures had not as yet been ascertained. Moreover the fact that some mutant types are repeatedly produced by others has become very prominent, indicating some kind of genetic relationship. Outward features are often observed to run parallel with such connections; in other cases they are so much alike, as to give another trustworthy basis for the assumption of closer relations.

Starting from these facts, we have tried to arrange the mutant types into distinct groups, analogous to those proposed by MORGAN and his disciples for Drosophila. Unfortunately, it is in our case mostly impossible to determine dihybrid splitting figures, such as constitute the basis for MORGAN'S deductions. Differences in the ratio of growth of the pollen tubes, preferential fertilization of one of the kinds of egg-cells, in those cases where there are two or more of them, deviations in the partial sterility of pollen-grains and of seeds are among the main causes which tend to discard numerical data from our present considerations. But the distinctness of our results does not seem to suffer from this difficulty. We start from the numbers of the chromosomes, and bring those mutants, which do not deviate in this respect from the main species, into our first group. In doing so, we find that almost all of them belong to the ordinary types of mutations. They are mostly due to recessive characters, as, for instance, the dwarfs and the brittle races.

In a large number of the other mutants one of the chromosomes is doubled, bringing the total number of these bodies up from seven to eight

in the haploid cells, and from fourteen to fifteen in the somatic nuclei. Here, however, it is clear that the relations are not so intimate as to allow the combining of all these forms into one single group. Evidently there are some main mutations, from which the others may be derived, or to which they can be subordinated. These primary mutations have been described by one of us under the name of dimorphic mutants (DE VRIES 1916), since, after self-fertilization, their progeny mainly consists for one part of individuals of the parental type, and for the remainder, of specimens of the type of the original species, *Oe. Lamarckiana*. All of them are heterogamous, since their characters are not reproduced by means of their pollen.

Of these dimorphic mutants there are six main types. Three are old ones, having been observed since the beginning of the experimental cultures of *Oe. Lamarckiana* (DE VRIES 1901, pp. 273, 288, 300). They are called mut. *lata*, mut. *scintillans* and mut. *spathulata*. The three others have been described for the first time in 1916, in the article already referred to. Their names are *Oe. cana*, *Oe. liquida* and *Oe. pallescens*.

Among these six forms two are known to produce almost annually and in a high percentage another definite type, which, however, is distinct in the two cases. Mut. *lata* produces about 2 percent *albida* and mut. *scintillans* throws off *oblonga* in a still larger percentage. In some instances it reached 6 percent of the whole progeny (DE VRIES 1913, p. 314). The reverse process does not occur, and therefore *albida* must be considered to be secondary to *lata*, and *oblonga* to be derived from *scintillans*.

Albida and *oblonga* are, however, not dimorphic in the sense given above. They constitute constant races, but this constancy is due in part to the presence of barren grains among their seeds. They have two kinds of female gametes but only one kind of these is represented among their male sexual cells. From this very distinct character, they may be called "one-and-one-half" or "sesquiplex" mutants.

In considering the remainder of the 15-chromosome mutants it is to be noted that one of them, *Oe. candicans*, is also a sesquiplex type, but closely related to *Oe. cana*, whereas almost all the others are to be derived either from *Oe. lata* or from *Oe. scintillans*.

Semi-gigas types are known to have 21 chromosomes, and their derivatives vary in this respect with numbers going from 15 to 20. Our race of *Oe. Lamarckiana gigas* has 28 of such rods, as has been discovered by Miss LUTZ and GATES. These special types, however, must be excluded from the present discussion since in them the changes are not limited to

a single chromosome, but include, at least in their origin, all of them at the same time.

Returning to the fourteen- and fifteen-chromosome mutants we are led to distinguish the following groups:

A. 14-chromosome mutants.
   1. *Nanella.*
   2. Forms with homogeneous pollen and ovules.
      a. With normal fibres; *blandina, decipiens.*
      b. With brittle fibres: *deserens, tarda, fragilis.*
   3. Sesquiplex mutants: ***simplex, secunda, compacta, elongata,*** *favilla, linearis.*

B. 15-chromosome mutants.
   1. *Lata* group.
      a. *Semi-lata.*
      b. Sesquiplex mutants: *albida, flava, delata.*
      c. *Subovata, sublinearis.*
   2. *Scintillans* group.
      a. Sesquiplex mutants: *oblonga, aurita, auricula, nitens, distans.*
      b. *Diluta, militaris, venusta.*
   3. *Cana* group: *candicans.*
   4. *Pallescens* group: *lactuca.*
   5. *Liquida.*
   6. *Spathulata.*

With the exception of those mentioned under B 2 b, the chromosomes for all of these types have been counted. The homogeneous mutants of the first section do not arise, as a rule, immediately from the parent species, but through the intermediate of half mutants (DE VRIES 1919). These are, in the sequence given above: *problandina, erythrina, rubrinervis, scindens* and *simplex.* All these have 14 chromosomes. To this group belongs the *rubricalyx* of GATES, which is a half mutant with 14 chromosomes. Analogous to *decipiens* is the *latifrons* which is a homozygous alethal segregate, derived from *rubricalyx* by SHULL (1921, page 362).

Of the sesquiplex mutants, enumerated in the first section, *Oe. simplex* is the original type, the others having sprung directly from it. In the group of *Oe. lata* the mutants *albida* and *subovata* are old forms, whereas *flava, delata* and *sublinearis* are new derivatives from the main type. From *scintillans* the forms called *militaris* and *venusta* have sprung, but *auricula* and *aurita* have originated independently. In their external

features, however, they closely resemble *Oe. oblonga*. *Oe. nitens* is so like *Oe. scintillans*, that it has been described originally under this name (DE VRIES 1918, p. 13) and *Oe. distans* and *Oe. diluta* have been derived from it.

Full descriptions of the stature and characters of the new types will soon be given in "Flora" and the "Zeitschrift für Botanik."

From this review we conclude that there are three large groups of mutants and probably four small ones. Among the first, one has only mutants with the normal number of chromosomes, and their characters are mostly of a recessive nature. It is probable that this group includes also the factors which determine the difference between the *laeta* and *velutina* gametes. We will call it the "central" group. This gives for the other divisions the name of "lateral" ones. Two of them are large, consisting of several externally similar or genetically related races. The four remaining groups are very small, each including as yet only one or two types.

In his classical researches on Drosophila MORGAN has compared the groups of the mutants of this fly with the relative size of the chromosomes and concludes that the larger rods carry the factors of the large associations, whereas the smaller groups are located in the smaller chromosomes. Now if we extend this principle to our case, we are led to expect three large chromosomes and four small ones in the haploid nuclei, or three large and four small pairs after copulation.

Unfortunately, the size of the chromosomes in *Oe. Lamarckiana* and its derivatives is so small, and the irregularities in their reduction phenomena are so prominent, that until now no direct answer can be given to our question.

But one of the Californian species, which are most intimately connected with *Oe. Lamarckiana*, gives better results. RALPH E. CLELAND (1922) has described the reduction divisions in the pollen mother cells of *Oe. franciscana* Bartlett. In mid-diakinesis the chromosomes are arranged in rings, each ring consisting of one or more pairs, as shown on CLELAND'S plate XXVII, figures 25–29. Differences in size are easily seen and most clearly in figure 28. Three pairs are large, but four are evidently smaller. The three first ones are on one side and the small ones on the other, constituting two sharp groups. In clearness and distinctness these preparations far surpass those obtained by previous authors from other species and from mutant races.

Therefore, it seems permissible to apply the results of CLELAND provisionally to our question. If we do so, our three large groups of mutants

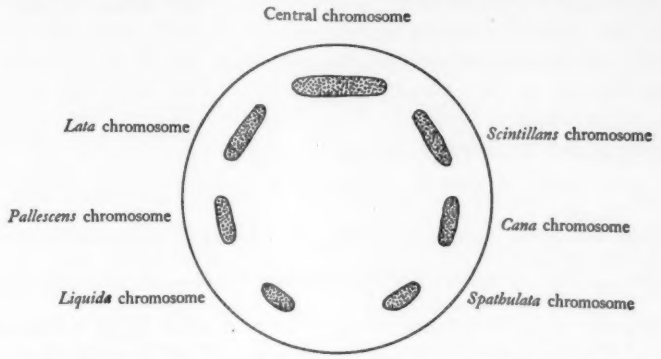


FIGURE 1.—Diagrammatic view of the seven haploid chromosomes in *Oenothera Lamarckiana*, to each of which, except the "central chromosome," is provisionally assigned one of the groups of 15-chromosome mutants.

obviously correspond, at least essentially, to the three large chromosomes and the four small groups to the four smaller chromosomes. We may then put the results of our investigation in the form of the following scheme (figure 1). In this scheme the position and relative size of the chromosomes have been drawn in accordance with figure 28 of CLELAND, but the axis of the figure has been turned a little so as to make it more clearly symmetrical. Of every pair only one rod has been reproduced.

Our central chromosome obviously corresponds to the chromosome I of SHULL (1921), which in the contrasting gametes of *Oe. Lamarckiana* carries the factors for the characters of *velutina* (I a) or those of *laeta* (I b). The new mutants *funifolia* and *pervirens* constitute also, according to the description of this author, members of the central group of mutant characters.

The lateral chromosomes have been given the names of their primary, dimorphic mutants, the sesquiplex and rarer mutants in each group being considered as of secondary importance.

Of course our scheme is only a provisional and in part an arbitrary one. But it is corroborated by the quite analogous results of BLAKESLEE with Datura. Many smaller arguments could be given to support it, as will be seen from the detailed description of our researches. We will here only adduce a single point, relating to the 21-chromosome types or *semi-gigas*. These proceed from *Oe. Lamarckiana*, with its two kinds of gametes, as

well as from *simplex*, all the sexual cells of which belong to the type of *laeta*. Now, whenever one of the lateral chromosomes mutates and is doubled, this does not seem to affect the sister chromosomes. But each time, when the central chromosome is doubled, all the lateral ones are affected in the same way, at the same time producing one or often a larger number of their possible mutations. "Very manifold groups of forms" (DE VRIES 1913 p. 333) may thereby arise, but their description must be put off till another opportunity.

LITERATURE CITED

BLAKESLEE, A. F., 1922 Variations in Datura, due to changes in chromosome number. Amer. Nat. **56**: 16-32.

BOEDIJN, K., 1920 Die Chromosomen von *Oenothera Lamarckiana* mut. *simplex*. Zeitschr. indukt. Abstamm. Vererb. **24**: 71-76.

CLELAND, R. E., 1922 The reduction divisions in the pollen mother cells of *Oenothera franciscana*. Amer. Jour. Bot. **9**: 391-413.

SHULL, G. H., 1921 Three new mutations in *Oenothera Lamarckiana*. Jour. Heredity **12**: 354-363.

VRIES, H. DE, 1901 Die Mutationstheorie. Vol. I, xii+648 pp. Leipzig: Veit & Co.

1913 Gruppenweise Artbildung. viii+365 pp. Berlin: Gebrüder Borntraeger.

1916 New dimorphic mutants of the Oenotheras. Bot. Gaz. **62**: 249-280.

1918 Kreuzungen von *Oenothera Lamarckiana* mut. *velutina*. Zeitschr. indukt. Abstamm. u. Vererb. **19**: 1-38.

1919 *Oenothera rubrinervis*, a half mutant. Bot. Gaz. **67**: 1-26.

THE THEORY OF PATH COEFFICIENTS

A REPLY TO NILES'S CRITICISM

SEWALL WRIGHT

Bureau of Animal Industry, U. S. Department of Agriculture, Washington, D. C.

Received November 21, 1922

TABLE OF CONTENTS

	PAGE
INTRODUCTION	239
General applications of the theory	240
Mathematics of path coefficients	243
Definitions	243
Path coefficients and correlation	244
Degree of determination	245
Precautions	246
Direction of influence	247
Application to causal relations	248
The relative importance of heredity and environment	250
The consequences of different systems of mating	251
Miscellaneous applications	252
Relation to multiple correlation	252
CONCLUSIONS	254
LITERATURE CITED	255

INTRODUCTION

The theory of path coefficients put forward by the present writer (WRIGHT 1921 a) as an aid in the biometric analysis of certain classes of data, is criticised by HENRY E. NILES (1922) in a recent number of this journal. The sweeping nature of his condemnation may be gathered from his conclusion:

"We therefore conclude that philosophically the basis of the method of path coefficients is faulty, while practically the results of applying it where it can be checked prove it to be wholly unreliable."

He begins with a statement of his conception of what the present writer claims as the purpose of the theory. This conception, we may say at once, is a complete misinterpretation.

"What occasions a result? What is its determining cause?

We have an answer to questions of this sort in many specific cases, but none of the attempts to produce a general formula universally applicable for

the solution of such questions has been entirely satisfactory. The present paper is a critical discussion of the latest solution offered, the method of 'path coefficients' as proposed by WRIGHT (1921 a)."

After a discussion of the nature of causation, he goes on:

"This method is claimed by WRIGHT (1921 a, b) to provide a measure of the influence of each cause upon the effect. Not only does it enable one to determine the effects of different systems of breeding, but provides a solution to the important problem of the relative influence of heredity and environment. To find flaws in a method that would be of such great value to science if only it were valid is certainly disappointing. The basic fallacy of the method appears to be the assumption that it is possible to set up *a priori* a comparatively simple graphic system which will truly represent the lines of action of several variables upon each other, and upon a common result."

This is followed immediately by a quotation from the present writer (WRIGHT 1921 a), apparently intended to corroborate NILES'S statement of the purpose of the theory.

"The method depends on the combination of knowledge of the degrees of correlation among variables in a system, with such knowledge as may be possessed of the causal relations. In cases in which the causal relations are uncertain the method can be used to find the logical consequences of any particular hypothesis in regard to them."

GENERAL APPLICATIONS OF THE THEORY

While the writer has always tried to choose his words with care, it is not impossible that an isolated sentence could be found in his papers on the subject which might seem to support NILES'S conception of his purpose. It is gratifying to find that the quotation actually given is such a satisfactory statement of it that any careful reader of NILES'S own paper would immediately see the wide difference between what NILES says of his purpose and what it really is.

The writer has never made the preposterous claim that the theory of path coefficients provides a general formula for the deduction of causal relations. He wishes to submit that the *combination* of knowledge of correlations with knowledge of causal relations, to obtain certain results, is a different thing from the *deduction* of causal relations from correlations implied by NILES'S statement. Prior knowledge of the causal relations is assumed as a prerequisite in the former case. Whether such knowledge is ever possible seems to be the subject of NILES'S philosophical discussion of the nature of causation. We will consider this question in more detail later, merely remarking here that to question it in the pragmatic sense intended by the writer is to question the utility of the so-called natural laws of the physical sciences, of physiology, genetics, of any field, in short,

in which it has been found possible to express relations between variable quantities in terms of mathematical formulae, exact or nearly so, within the limits of human observation.

The writer also wishes to submit that "finding the logical consequences" of a hypothesis in regard to the causal relations does not depend on any prior assumption that the hypothesis is correct. Neither does it imply that the theory of path coefficients by itself gives a method of proving such a hypothesis. It does, of course, follow that if one of the logical consequences of a hypothesis is absurd the hypothesis is untenable and must be modified; on the other hand, if the logical consequences can be shown to agree with independently obtained results it contributes to the demonstration of the truth of the hypothesis in the only sense which can be ascribed to the truth of a natural law. As NILES says (italics his):

> "In no case has it been proved that there is an inherent necessity in the laws of nature. . . .
> It seems clear that perfect correlation *when based upon sufficient experience,* is causation in the scientific sense."[1]

The writer would substitute JOHN STUART MILL'S expression "uniform antecedence" for "perfect correlation," but accepts the viewpoint that our conceptions of causation are based merely on experience.[1]

We may divide the applications of the theory into three cases: (1) Where the causal relations among the variables may be considered as known; (2) where enough is known to warrant an hypothesis or alternative hypotheses; and (3) where even an hypothesis does not seem justified.

In case (1) we have intimated above that something can be done by combining the two kinds of knowledge, that of causation and that of correlation. This something consists of measurements of the relative importance of the influence along various paths in the case in question and predictions of the effect which control of one or more of the factors will have on the variations of other factors, their correlations, etc. In case (2) we have stated that the logical consequences of a hypothesis as to causal relations, can be worked out, permitting comparison with new facts. As to case (3), it might seem that nothing could be done. We have nothing to combine with our knowledge of the correlations.

[1] As in many other cases it is difficult to be certain of representing NILES's viewpoint correctly. He devotes much space to the development of the idea that causation implies no "inherently necessary connection between things" but makes the following statement later:

> " . . . The closeness of agreement between calculated or expected and observed values is an unscientific criterion by which to judge the validity of such a system" i.e., a system of causal relations.
> It would be interesting to know what scientific criterion he had in mind.

If, however, we go ahead in the only way that can possibly be justified, we arrive at certain valuable, but not novel, expressions for the correlation between the given variable and the whole group of others and also for the particular linear function of the others which gives this correlation. This gives a means of estimating the most probable value of the given variable for known values of the others. As to interpretation in terms of causation, NILES makes the following quotation from the writer, which may be repeated.

"One should not attempt to apply in general a causal interpretation to solutions by the direct methods. In these cases, determination can usually be used only in the sense in which it can be said that knowledge of the effect determines the probable value of the cause. This is the sense in which PEARSON's formula for multiple regression must be interpreted."

The direct methods referred to above consist of a group of formulae derived from the theory of path coefficients, which are applicable only in the symmetrical system of relations which must be used where nothing is known of causal relations. It was not recognized at the time these were worked out, but was recognized in time for insertion and discussion in the original paper on the subject, that these formulae were none other than PEARSON'S formulae for multiple correlation and multiple regression, expressed in a different (and much less elegant) mathematical form.

A few lines below NILES gives the following opinion of the applicability to the determination of causal relations in this case in which we know least.

"Statistical methods, particularly multiple correlation, indicate causes when they are used with common sense and upon the data of critical experiments."

Comment is unnecessary. He goes on,

"But the method of path coefficients does not aid us because of the following three fallacies that appear to vitiate this theory. These are (1) the assumption that a correct system of the action of the variables upon each other can be set up from *a priori* knowledge; (2) the idea that causation implies an inherently necessary connection between things, or that in some other way it differs from correlation; (3) the necessity of breaking off the chain of causes at some comparatively near finite point."

The writer cheerfully accepts (1) and at least the first part of (2) as false *general* assumptions. Their falsity does not vitiate a proper application of the theory. As to (3) it expresses a limitation on analysis in particular cases but not in all cases. We will come back to these points later. So much for the present as to NILES'S general criticisms.

MATHEMATICS OF PATH COEFFICIENTS

Let us get to the crux of the matter. Is the theory valid as a purely mathematical proposition? If it is not, misconceptions of it are supremely unimportant. If it is valid, the validity of the suggested applications can hardly but be granted.

On this question, NILES has little to say and is apparently uncertain. He says of the general formula, which, with the definition of a path coefficient, includes the whole theory:

> "The pure mathematics by which this is shown is apparently faultless in the sense of mere algebraic manipulation, but it is based upon assumptions which are wholly without warrant from the standpoint of concrete, phenomenal actuality."

In connection with the definition of a path coefficient, he does, however, say that the writer has overlooked a simple point of logic. It is easy to show that what he has reference to is nothing more than a possible lack of generality in the original form of statement of the definition, easily corrected and which in no wise invalidates the theory. The preferable form was indeed given later in the same paper.

DEFINITIONS

We will start with a number of variable quantities, certain of which are linear functions of others. Figure 1 is meant to represent such a system in which the arrows indicate which quantities are combined. Each value of A for example is supposed to be derived by adding particular values of D and E, or multiples of these values. The system may be of any degree of complexity.

As figure 1 is drawn, factors A, B and C, contribute to the variability of X. The standard deviation of X due to A, $\sigma_{X \cdot A}$, was defined as follows:

> "We will start with the assumption that the direct influence along a given path can be measured by the standard deviation remaining in the effect after all other possible paths of influence are eliminated, while variation of the causes back of the given path is kept as great as ever, regardless of their relations to the other variables which have been made constant. Let X be the dependent variable or effect and A the independent variable or cause. The expression $\sigma_{X \cdot A}$ will be used for the standard deviation of X, which is found under the foregoing conditions, and may be read as the standard deviation of X due to A."

If we make all of the immediate factors except A constant, the variation left in X, measured by ${}_{CB}\sigma_X$ (indicating constant factors by subscripts to the left), must be due wholly to the direct influence of A, i.e., each value of X is a certain multiple of A. This direct influence is exerted, however, in a population in which the variation of A itself is reduced because of its correlations with B and C. In order to measure the variation of X relative to the entire direct influence of A in the original population, the expression for the standard deviation of X for constant B and C (${}_{CB}\sigma_X$) must be multiplied by the ratio of the original standard deviation of A to its value in the population in which B and C are constant.

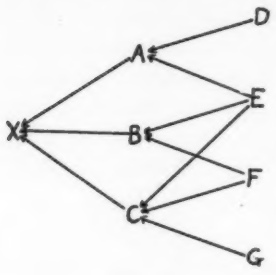


FIGURE 1

Thus,
$$\sigma_{X.A} = {}_{CB}\sigma_X \frac{\sigma_A}{{}_{CB}\sigma_A}$$

It may be admitted that the operations suggested by the verbal definitions could not be literally carried out in extreme cases and the definition is therefore imperfect. The above formula, however, which was given later in the paper can always be calculated.

The path coefficient, $p_{X.A}$, is the ratio of this standard deviation of X due to A to the total standard deviation of X.

$$p_{X.A} = \frac{\sigma_{X.A}}{\sigma_X} = \frac{{}_{CB}\sigma_X}{{}_{CB}\sigma_A} \frac{\sigma_A}{\sigma_X}$$

Path regression, the direct effect of a unit variation in A on variation in X in actual units, is given by the formula,

$$p\ reg_{X.A} = p_{X.A} \frac{\sigma_X}{\sigma_A} = \frac{\sigma_{X.A}}{\sigma_A} = \frac{{}_{CB}\sigma_X}{{}_{CB}\sigma_A}$$

PATH COEFFICIENTS AND CORRELATION

Two variables are in general correlated with each other if they are determined in part by one or more common factors. The product of the path coefficients in a chain of paths which connects them through such a common factor measures the contribution of the chain in question to their correlation. It was shown in the original paper that the sum of the contributions of all such independent connecting chains equals the coefficient of correlation.

In figure 2, X and Y are determined in part by two common factors B and C and it is supposed that B and C are themselves correlated through common factors. According to the principle just stated, the products of the coefficients along any chains by which B and C may be connected are summed up in the coefficient of correlation r_{BC}. X and Y are connected by four independent chains, X-B-Y, X-C-Y, X-B-C-Y and X-C-B-Y. Summing up the products of the path coefficients (indicated by small letters) along these paths, and using r_{BC} to represent those connecting B and C, we have

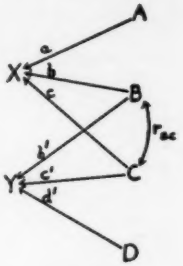


FIGURE 2

$$(1) \qquad r_{XY} = b\,b' + c\,c' + b\,r_{BC}\,c' + c\,r_{BC}\,b'$$

The formula is perfectly general since more complex systems can always be reduced by stages to a system like that in figure 2. It can be written in the following form which is frequently more convenient.

$$r_{XY} = a\,r_{AY} + b\,r_{BY} + c\,r_{CY}$$

DEGREE OF DETERMINATION

A second property of path coefficients was demonstrated. If one variable is completely determined by a number of others, the sum of the squares of the path coefficients leading to it, plus certain terms expressing joint determination by correlated variables, equals unity. A joint term of this kind is twice the product of the two path coefficients times the coefficient of correlation between the two variables in question. We have then:

$$(2) \qquad a^2 + b^2 + c^2 + 2bc\,r_{BC} = 1$$

Because of this property the squares of the path coefficients give a useful measure of the degree of determination. Each one measures the portion of the squared standard deviation for which the factor in question is responsible.

If we try to express the correlation of a variable (X) with itself in terms of path coefficients according to equation 1 we get equation 2. We also see that equation 2 can be expressed in the following simple form which can easily be generalized—

$$a\,r_{AX} + b\,r_{BX} + c\,r_{CX} = 1$$

It is evident that equation 2 is merely a limiting form of equation 1.

The definition of a path coefficient including its relation to standard deviation, its relation to correlation expressed in equation 1 and to determination expressed in equation 2 are essentially all there is to the theory of path coefficients.

PRECAUTIONS

In applying these formulae there are certain precautions which it may be well to mention. In equating a coefficient of correlation to the contribution of the chains of paths which connect the two variables in question, (equation 1) care must be taken to avoid duplication of chains. In figure 3, M, N and B are common factors of X and Y. The chains X-M-Y, X-N-Y and X-B-Y connect X and Y. They are not however independent chains. M and N are common factors of X and Y only because B is. The chain X-B-Y sums up all causes of correlation between X and Y which are indicated in the figure.

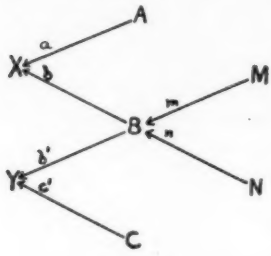


FIGURE 3

Thus $r_{XY} = b\,b'$

If we do not take B into consideration, it is legitimate to go back to the more remote common factors in tracing connecting chains. Since $p_{XM} = b\,m$, $p_{YM} = b'\,m$ by equation 1, we have

$$\begin{aligned} r_{XY} &= bm\ b'm + bn\ b'n \\ &= bb'\ (m^2 + n^2) \\ &= bb' \text{ (by equation 2, giving the same result as before).} \end{aligned}$$

With this precaution in mind there should be no difficulty in determining which chains of paths contribute independently to the correlation between two variables and which merely analyze in more detail, the contribution of another chain.

A similar precaution is, of course, necessary in expressing complete determination of one variable by others. We can express the complete determination of X by A, M and N ($a^2 + b^2m^2 + b^2n^2 = 1$), but if we take B into consideration we can only consider the contributions of M and N for which B is not an intermediary. As the diagram shows no such contributions, we have simply $a^2 + b^2 = 1$.

It should hardly be necessary to point out that chains of paths which indicate the coöperation of two variables in determining a third, do not contribute to the correlation between the two former. In a diagram in which the direction of determination along each path is indicated by an arrow, it is legitimate to follow the arrows directly from one variable to the other, or to go back along the arrows from one to a common factor and then forward to the other, but never to go forward and then back, in

listing the connecting chains which contribute to the correlation. Certain of NILES's criticisms are invalidated by his failure to apply the method correctly in this respect. He says,

"Perhaps in tracing the chains of causes we are not allowed to come to the effect through something that follows it. Such a rule would be meaningless when the true relation of cause and effect, that of invariable association, is kept in mind."

DIRECTION OF INFLUENCE

Mathematically, a path coefficient is not a symmetrical relation between variables like a coefficient of correlation. Direction must be indicated in some way. This does not mean that either of two variables may not be considered as determining the other but merely that the reverse paths, $p_{X.A}$ and $p_{A.X}$ are not in general equal. This can easily be seen from the definition. It can also be illustrated by a simple example. Assume that each value of X (figure 4) is the sum of values of A and B which are independent variables ($r_{AB}=0$), with equal standard deviations. X is clearly determined equally by A and B, i.e. $a^2=b^2$.

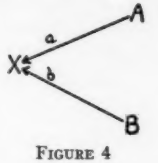


FIGURE 4

From 2, $$a^2+b^2=1$$
$$a=b=\sqrt{1/2}$$
From 1, $$r_{XA}=r_{XB}=\sqrt{1/2}$$

We can just as legitimately turn the diagram around (figure 5) and consider A as determined by X and B. It is their difference. We still have the same correlations, of course.

From 1, $$r_{AX}=x+yr_{XB}=\sqrt{1/2}$$
$$r_{AB}=xr_{XB}+y=0$$

Solving, we obtain, $x=\sqrt{2}$ and $y=-1$. These values could have been obtained from fairly obvious *a priori* considerations. The important point is that $p_{X.A}=a=\sqrt{1/2}$ is widely different from $p_{A.X}=x=\sqrt{2}$.

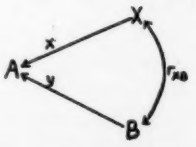


FIGURE 5

This illustration also brings out two other interesting points. From the value of x we see that a path coefficient, unlike a coefficient of correlation, may take values greater than unity. Secondly, the fact that $p_{A.B}=y=-1$, may seem surprising to the reader in view of the assumption that A and B are independent variables.

A coefficient of correlation gives an absolute measure of the correlation between two variables in a given body of data. No such absolute interpretation can be put on a path coefficient. It measures the influence of one variable on another from a particular viewpoint, that indicated in

the diagram of relations. That this is so is obvious from the definition. It is a consideration which makes it necessary to use a great deal of care in the interpretation of applications of the method. We do not hesitate to say that it is a dangerous method to use without a thorough understanding. This, however, does not invalidate its proper use.

Returning to figure 5, we see that A is necessarily completely determined by subtracting B from X. Does equation 2 still hold? Making the substitutions we find in fact that it does hold $x^2+y^2+2\,xy\,r_{XB}=1$. $(\sqrt{2})^2+(-1)^2+2\sqrt{2}\,(-1)\,\sqrt{1/2}=1$.

Suppose that we combine the two systems of figures 4 and 5, i.e., first add B to A to get X and then subtract B from X giving us A again. Here A appears as a "cause" of itself. Figure 6 shows the diagram of relations. We must indicate in it, a correlation of $+1$, between B and B' in order to show that it actually is B that is being subtracted from X. We must not, however, indicate any additional path connecting A and A' because their identity is a result beyond our choice which should come out as a consequence of the other relations. Substituting the values of the various path coefficients as already found, we have:

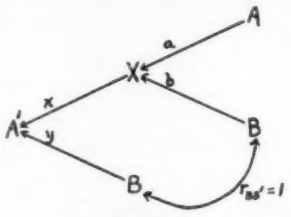


FIGURE 6

$$r_{A'X}=x+y\,b=\sqrt{1/2}=r_{\ X}$$
$$r_{A'B}=bx+y=0=r_{AB}$$
$$r_{A'A}=ax=1$$

A' is thus indistinguishable from A in its correlations as should be the case.

In general, as a purely mathematical proposition, we may look upon either one of two variables as a factor contributing toward the determination of the other. Indeed, as just seen, the opposite viewpoints may be taken in different places in the same system of relations. But whatever viewpoint is taken of a particular path, it must be held consistently. NILES's failure to grasp this point leads him to omit arrowheads or any other indications of direction of influence from his diagrams and, of course, leads to "absurd results" when he introduces a given factor into the system twice.

APPLICATION TO CAUSAL RELATIONS

In applying the theory to concrete variables, consistent results can be obtained in whatever way the arrows are drawn. The writer has not

hesitated to calculate coefficients for reverse paths (WRIGHT 1921 b, d) The interpretation, of course, depends on the direction in which the path is drawn. In dealing with causal relations, the interpretation is usually more illuminating if the arrows are drawn so as to indicate the direction of causation, i.e., from cause to effect.

If the writer understands him correctly, NILES seems to hold that the concept causation does not involve direction and for this reason objects to direction in path coefficients. This point of view seems to be indicated in the last quotation given (page 247) and in his reiteration of the statement that "causation is correlation."

Nevertheless, he quotes with apparent approval from PEARSON and indirectly from JOHN STUART MILL:

> "That a certain sequence has occurred and recurred in the past is a matter of experience to which we give expression in the concept causation." . . .
> " 'Causation' says JOHN STUART MILL, 'is uniform antecedence' and this definition is perfectly in accord with the scientific concept."

Surely NILES does not wish us to understand that it is a matter of indifference whether the earlier or the later event in a sequence is considered a cause of the other.

The real difficulty is in distinguishing a relation of cause and effect from a correlation due to common causes. Causation implies not merely that the effect follows the cause but also that an adequate succession of events, continuous in time and space, intervenes between them.

Pure causation implies that there is no event or condition earlier than both the cause and effect in question to which both trace through independent sequences of events.

In representing a system of relations, causation *must* be presented by a directed path as in case 1, figure 7. Correlation through a known common cause is of course to be represented as in case 2. Correlation through the *totality of unknown causes* can be represented as in case 3. A relation of causation complicated by correlation through unknown common causes can be represented as in case 4.

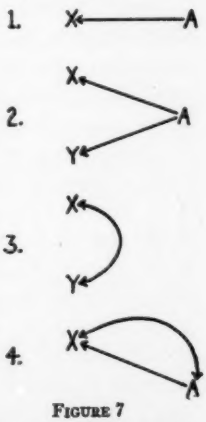


FIGURE 7

Can we ever know the causal relations with sufficient completeness to make it possible to make a satisfactory analysis by the method of path coefficients? NILES denies this. He quotes PEARSON as follows:

"The causes of any *individual* thing thus widen out into the unmanageable history of the universe. The causes of any finite portions of the universe lead us irresistibly to the history of the universe as a whole."

This looks discouraging. PEARSON, however, is here talking about the total cause. In practice we find that we can satisfactorily isolate a portion of the universe and deal with causation relative to this limited system. To illustrate, the genetic constitutions of two guinea-pigs, chosen at random from a certain stock, undoubtedly trace back to numerous common sources. They must be closely correlated, *but this is relative to the whole universe of things*. If we are dealing only with relations within this stock of guinea-pigs, they are uncorrelated from the mere fact that they were chosen at random. We have taken the average amount of relationship as the zero of relationship within this stock.

Stated in another way, science does not deal with total causes. It deals with the causes of differences which are the same as the differences of the total causes. In subtracting the total cause of one event from another there is an enormous cancellation of common factors.

THE RELATIVE IMPORTANCE OF HEREDITY AND ENVIRONMENT

Let us illustrate by another biological example. The totality of causes determining the characteristics of an animal or plant is usually divided into two great categories, heredity and environment. The problem as to the relative importance of heredity and environment in determining the characteristics of a single given individual has no meaning. Both are absolutely essential. You can not raise a duck from a hen's egg nor can you raise a chicken from a hen's egg in the environment of a fiery furnace. On the other hand, the problem as to their relative importance in determining variation, or *differences*, within a given stock, has a perfectly definite meaning and can often be solved with great ease. If we raise the birds from a mixed lot of eggs from hens, ducks and turkeys, under as nearly constant conditions as we are able, we will find great differences in the results which can safely be ascribed nearly 100 percent to heredity. This may even be true of birds from eggs all laid by one mongrel hen. If, on the other hand, we take a strain of fowls that has been inbred for years and selected for conformity to a certain type, but raise the chicks with different quantities and kinds of feed, we are likely to find much variation which must be attributed nearly 100 percent to environment. The percentage determination by each factor in intermediate cases has a perfectly definite meaning. The writer, for example, (WRIGHT 1920), found by analysis by the method of path coefficients that the variation

in color pattern in a certain stock of guinea-pigs was determined 42 percent by heredity and 58 percent by environment.

This meant, according to the definition, that 42 percent of the squared standard deviation of this stock should disappear, if the variation in heredity could be eliminated. Such a stock, descended from a single mating in the twelfth generation of brother-sister mating, happened to be on hand. The same method of analysis indicated only 3 percent determination by heredity and 97 percent by environment. The square of its standard deviation, a quantity not used in the analysis, was 43 percent less than that in the random-bred stock, in excellent agreement with expectation.

THE CONSEQUENCES OF DIFFERENT SYSTEMS OF MATING

The most satisfactory applications are naturally those in which a complete system of causal relations is at hand. The most extensive application which the writer has yet made (WRIGHT 1921 b, c, d, 1922), has been one of this kind in which the Mendelian theory of heredity furnished the system of relations. Certain of the consequences of various systems of mating were deduced by the use of path coefficients. NILES dismisses the validity of the results in a paragraph already quoted, on the grounds of a "basic fallacy" in the "assumption that it is possible to set up *a priori* a comparatively simple graphic system which will truly represent the lines of action of several variables upon each other and upon a common result." This seems to mean that NILES considers the Mendelian theory to be so indefinite or so complex that it is not legitimate to analyze its consequences by mathematical methods. He is thus attacking the application of ordinary algebraic methods as by JENNINGS, FISH, PEARL, WENTWORTH and REMICK, and others (cf. WRIGHT 1921 b), in dealing with this same problem just as much as he is attacking the use of the method of path coefficients.

It may be pointed out, incidentally, as evidence of the validity of the method of path coefficients that it has given results identical with those obtained by the algebraic methods as far as the latter have been carried.

It was of course realized that the "concrete, phenomenal actuality" of the results was not proved by the analysis by path coefficients. This rests on the validity of the premises, i.e., on the evidence for Mendelian heredity. The paper began with a quotation from EAST and JONES on the universality of Mendelian inheritance under sexual reproduction, as the justification for the analysis.

MISCELLANEOUS APPLICATIONS

The application to cases in which the causal relations have less support than the Mendelian hypothesis is of course just as legitimate, as long as the results are properly interpreted. NILES discusses at considerable length an analysis of an hypothesis in regard to the factors which determine the birth-weight of guinea-pigs. How far this hypothesis completely represents the causal relations need not be discussed here. Suffice it to say that much can be said in justification of each path that is represented. In fact, it represents graphically a combination of two hypotheses discussed by MINOT (1891). A combination of the methods of correlation and path coefficients enables one to make a more complete analysis of the consequences of these hypotheses than could Professor MINOT with the purely verbal logic then available and leads to different conclusions from those which he drew. The "absurd results" which NILES obtains by ignoring the direction of the paths of influence have already been referred to.

Following discussion of this case, NILES tests the theory by applying MINOT's guinea-pig hypothesis to a group of correlations relating to the weight of bean seeds and to another group relating to the heat produced in basal metabolism. He obtains "ridiculous" results. There are two alternative possibilities. One might conclude (1) that MINOT's guinea-pig hypothesis does not happen to apply in these cases, or, (2) that it does and that hence the mathematical method by which ridiculous consequences were brought to light must have been at fault. NILES takes the latter course.

The writer does not feel called upon to suggest systems of relations to which it would be worth while to apply the method of path coefficients in these cases. Such systems are hypotheses. The formulation of hypotheses is emphatically the business of one who is thoroughly familiar with the realities of the case.

RELATION TO MULTIPLE CORRELATION

Let us take up again the case in which nothing is assumed as to causal relations in a system of variables X A B C and attempt to discover how far X is determined, in a purely mathematical sense, by the other variables (figure 8). Let O represent residual factors. The only way in which A, B and C can be related symmetrically is to connect them by double-headed arrows indicating that all connecting chains of paths are summed up in coefficients of correlation.

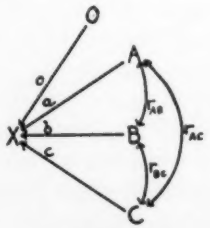


FIGURE 8

Applying equation 1 we have

$$r_{XA} = a + b\, r_{AB} + c\, r_{AC}$$
$$r_{XB} = a\, r_{AB} + b + c\, r_{BC}$$
$$r_{XC} = a\, r_{AC} + b\, r_{BC} + c$$

On solving these simultaneous equations we obtain values of the path coefficients a, b and c. The most probable deviation of X from its mean value (X') for known deviations of A, B and C (A', B' and C') can be found at once by adding the formulae for path regression.

$$\frac{X'}{\sigma_X} = a\frac{A'}{\sigma_A} + b\frac{B'}{\sigma_B} + c\frac{C'}{\sigma_C}$$

The total extent to which X is determined by A, B and C, can be calculated from equation 2.

$$a^2 + b^2 + c^2 + 2ab\, r_{AB} + 2\, ac\, r_{AC} + 2\, bc\, r_{BC} + o^2 = 1$$
$$1 - o^2 = a\, r_{AX} + b\, r_{BX} + c\, r_{CX}$$

If one wishes, one can solve the above group of linear simultaneous equations with the help of determinants.

$$a = \frac{\begin{vmatrix} r_{XA} & r_{BA} & r_{CA} \\ r_{XB} & 1 & r_{CB} \\ r_{XC} & r_{BC} & 1 \end{vmatrix}}{\begin{vmatrix} 1 & r_{BA} & r_{CA} \\ r_{AB} & 1 & r_{CB} \\ r_{AC} & r_{BC} & 1 \end{vmatrix}}$$

The expressions for b and c are analogous. If we let

$$\Delta = \begin{vmatrix} 1 & r_{AX} & r_{BX} & r_{CX} \\ r_{XA} & 1 & r_{BA} & r_{CA} \\ r_{XB} & r_{AB} & 1 & r_{CB} \\ r_{XC} & r_{AC} & r_{BC} & 1 \end{vmatrix}$$

and let Δ_{XA} represent the minor made by deleting column X, row A, with its appropriate sign, we can express the path coefficients a, b and c in more compact form.

$$a = -\frac{\Delta_{XA}}{\Delta_{XX}},\ b = -\frac{\Delta_{XB}}{\Delta_{XX}},\ c = -\frac{\Delta_{XC}}{\Delta_{XX}}$$

This gives us at once PEARSON'S expressions for multiple regression[2] and correlation.

[2] The minus sign before the bracket was omitted in giving this formula in the original paper. The formula was, however, correctly applied in the illustrations given.

$$\frac{X'}{\sigma_X} = -\left(\frac{\Delta_{XA}}{\Delta_{XX}}\frac{A'}{\sigma_A} + \frac{\Delta_{XB}}{\Delta_{XX}}\frac{B'}{\sigma_B} + \frac{\Delta_{XC}}{\Delta_{XX}}\frac{C'}{\sigma_C}\right)$$

$$R_{X \cdot ABC} = \sqrt{1 - o^2} = \sqrt{1 - \frac{\Delta}{\Delta_{XX}}} \quad \text{by an easy transformation.}$$

These formulae can obviously be generalized. That a simple demonstration of PEARSON'S formulae should drop out of the application of the theory of path coefficients to a symmetrical system of relations would be a remarkable coincidence if the theory were not mathematically sound.

CONCLUSIONS

NILES'S condemnation of the method of path coefficients on the grounds that "philosophically the basis is faulty" and that "practically the results of applying it where it can be checked prove to be wholly unreliable" we have shown, we believe, to be without sound foundation.

His first criticism is vitiated throughout by the fallacy that because the method does not enable one to accomplish certain things, (which, it happens, were never claimed for it) it is of no possible use. The failure to consider specifically the applications of the method, which have actually been suggested and made, seems to be based on what appears to the writer an untenable attitude toward causation and laws of nature as working scientific concepts.

His second general criticism is based on alleged "ridiculous" consequences of the method. These are obtained, however by incorrect mathematics, apparently consequent upon a failure to recognize that a path coefficient is not a symmetrical function of two variables, but that it necessarily has direction. He, himself, can find no fault with the algebraic processes by which the fundamental formulae were derived. The validity of these formulae is checked, it may be added, by the fact that they give results identical with those obtained by ordinary algebraic methods in analyzing the consequences of Mendelian heredity in fairly complicated systems of mating, and by the fact that very simple demonstrations of PEARSON'S formulae for multiple correlation and regression can be obtained from them.

The method of path coefficients does not furnish general formulae for deducing causal relations from knowledge of correlations and has never been claimed to do so. It does, however, within certain limitations, give a method of working out the logical consequences of a hypothesis as to the causal relations in a system of correlated variables. The results are obtained by a combination of the knowledge of the correlations with

whatever knowledge may be possessed, or whatever hypothesis it is desired to test, as to causal relations. Such results may contribute toward the analysis of the causal relations by giving a basis for comparison with independently obtained results. A disagreement necessitates modification of the hypothesis while agreement contributes toward the demonstration of its truth, in the only sense in which truth can be ascribed to a scientific hypothesis.

Summing up, the criticisms offered by NILES in no way invalidate the theory of path coefficients or proper applications of it to statistical problems.

LITERATURE CITED

MINOT, C. S., 1891 Senescence and rejuvenation. Jour. Physiol. **12**: 97–153.

NILES, H. E., 1922 Correlation, causation and Wright's theory of "path coefficients." Genetics **7**: 258–273.

WRIGHT, SEWALL, 1920 The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs. Proc. Nation. Acad. Sci. **6**: 320–332.

1921 a Correlation and causation. Jour. Agric. Res. **20**: 557–585.

1921 b Systems of mating. I. The biometric relations between parent and offspring. Genetics **6**: 111–123.

1921 c Systems of mating. II. The effects of inbreeding on the genetic composition of a population. Genetics **6**: 124-143.

1921 d Systems of mating. III. Assortative mating based on somatic resemblance. Genetics **6**: 144–161.

1922 Coefficients of inbreeding and relationship. Amer. Nat. **56**: 330–338.

POSTSCRIPT.—The opportunity has courteously been given the writer of seeing Mr. NILES's counter-reply. As it appears to him that an adequate answer to all of the points which NILES raises may be found on careful reading of his present paper, he is willing to rest his case at this point.

THE METHOD OF PATH COEFFICIENTS[1]

AN ANSWER TO WRIGHT

HENRY E. NILES
2010 Edgewood St., Baltimore, Maryland

Received January 2, 1923

It seems to me that my condemnation of the method of path coefficients on the grounds that "philosophically the basis is faulty" and that "practically the results of applying it where it can be checked prove to be wholly unreliable" was a just condemnation and one that is not invalidated by Doctor WRIGHT's answer to my paper. Most of his answer is beside the point. He spends comparatively little time dealing with the philosophic basis of his theory. He states that I made mistakes in applying his method to the two test problems I gave, but he does not show wherein my mistakes lay. The method by itself did not show that my incorrect results were wrong. He has not shown his method to be superior to the easily interpreted method of multiple correlation.

The greatest point of dispute between WRIGHT and myself seems to be in regard to the nature of causation and correlation. WRIGHT firmly maintains his old position and I am unconvinced that I was in error, except in one unimportant detail. I have never attacked the mathematics of the method of "path coefficients" because it seems sound enough *when the preliminary assumptions regarding the basis of the method are granted*, but I do not grant them.

In my paper (NILES 1922, p. 262) I claimed that WRIGHT's theory was vitiated by the following three things:

"(1) The assumption that a correct system of the action of the variables upon each other can be set up from *a priori* knowledge;
(2) The idea that causation implies an inherently necessary connection between things, or that in some other way it differs from correlation;
(3) The necessity of breaking off the chain of causes at some comparatively near finite point."

In regard to point (1), WRIGHT claims that if ridiculous results are obtained from any system that then the system and the hypothesis

[1] Papers from the Department of Biometry and Vital Statistics, School of Hygiene and Public Health, JOHNS HOPKINS UNIVERSITY, No. 88.

behind it must be wrong, but the mathematical theory remains valid. He does not, and, I think, can never, establish a generally applicable test of the validity of any system or hypothesis set up except the agreement between the results given by it and results expected on the basis of other calculations, judgments, or observations. The test applications of the method of path coefficients which I gave in my original paper show that it is possible to obtain by this method an apparently consistent, reasonable value for a correlation coefficient although when compared with the coefficient calculated directly from the data there is a great difference between the two values. This is not due to my disregarding the direction of the path coefficients. A diagram of causes similar to one of WRIGHT'S was set up and similar equations were used because I feared that I had not mastered the rules of the game sufficiently to be confident of deriving the correct equations for a system of causes not given by WRIGHT. The omission of arrow heads in the diagrams in my paper was a draughtsman's error. In each diagram the direction of the path coefficient is from right to left, and all work was done with this in mind. When such a disagreement as that just mentioned occurs, WRIGHT would probably say that the diagram of causes was incorrect and that another need be tried. But the point is that in the case cited WRIGHT'S method alone would not have shown this. We had to know the true value of the correlation to find that its theoretical value was wrong. If the true correlation coefficients can be calculated from the raw data it seems to me to be useless, to put it mildly, to set up a system of path coefficients. But unless the true value can be calculated there seems to be little basis for accepting the value given by the method of path coefficients.

The writer is glad to have this chance to correct his statement that the "closeness of agreement between calculated or expected and observed values is an unscientific criterion by which to judge the validity of such a system;" i.e., of path coefficients. This taken by itself is indeed foolish. One must test the results of hypotheses with the concrete observed actualities. In its context the statement was not such a mistake. What was meant was that the method of path coefficients could not be justified merely by the fact that *at times* it gave results in agreement with observed realities, or results consistent with what we only think to be the case. When it is realized that the philosophical basis of the system is unsound, the fact that it can at times be made to give results consistent with reality will not be considered proof of the validity of the theory.

WRIGHT admits that the first part of point (2) "the idea that causation implies an inherently necessary connection between, things," is a false

assumption, but he claims that its falsity does not vitiate a proper application of his theory. He does not seem ready to agree that causation does not differ from very high or perfect correlation in a very large sample of data. We may find a very high degree of correlation between two variables in a set of data and yet we may be sure that one variable is not the cause of the other; but we are sure only because we know from *experience not included in the data upon which are based our correlation* that if we had taken into account this greater experience that the correlation would not have been very high. It would be justifiable, to say that *in the sample studied* the value of one of the variables, the first in time, *caused* the particular values of the other to occur. This is correct but it is confusing and it seems best to say that in the sample there is correlation but that we know from knowledge not considered in our calculations that the high value of the coefficient of correlation does not hold for the whole of our experience.

An example of correlation in a sample without causation throughout our range of experience may help to make the matter clearer. If a man plants a tree in his front yard when his son is born he may notice for twenty years the high correlation between the growth of the tree and of his son and yet he will not believe that the growth of one is due to the growth of the other or that both are due to a common cause. He knows that the tree he planted is behaving as trees have been observed to behave since time immemorial and that his son is behaving as sons have always been observed to behave. He can be sure that the correlation will not continue when his son becomes a man and stops growing. No one has yet proved that it is *necessary* for the son to stop growing; but all sons, so far as we know, have in the past stopped growing at some comparatively early age, and so we say we are sure that this one will stop growing.

Another example is furnished by Greek natural science. The Greeks observed that tadpoles appeared in stagnant puddles where vegetation was decaying, so they said that tadpoles were caused by decaying matter changing into this form of life. Wherever they observed stagnant water they found tadpoles and they didn't find them elsewhere. It was therefore *scientifically sound* for them to consider stagnant water the cause of tadpoles. But we know that stagnant water does not cause tadpoles since *by further observation* we have found that the highest correlation is not between stagnant water and tadpoles but between the appearance of queer jelly-like masses in stagnant water and later of tadpoles. These jelly-like masses were found by observation to be deposited by frogs and now we have a much more detailed knowledge than the ancients had of what causes the tadpoles. Our knowledge is more complete than was that

of the Greeks, but we can not be sure that we have reached the final explanation. Both the Greeks and ourselves have observed correlations and called them instances of causation. Perfect or complete correlation, when based upon a sufficiently large sample, is causation in the scientific sense.

The following quotation is of interest in showing WRIGHT's attitude toward correlation and causation (italics are WRIGHT's) (see WRIGHT 1923, pages 240, 241):

"He (i.e., WRIGHT) wishes to submit that the *combination* of knowledge of correlations with knowledge of causal relations, to obtain certain results, is a different thing from the *deduction* of causal relations from correlations implied by NILES's statement. Prior knowledge of the causal relations is assumed as a prerequisite in the former case. Whether such knowledge is ever possible seems to be the subject of NILES's philosophical discussion of the nature of causation. We will consider this question in more detail later, merely remarking here that to question it in the pragmatic sense intended by the writer is to question the utility of the so-called natural laws of the physical sciences, of physiology, genetics, of any field, in short, in which it has been found possible to express relations between variable quantities in terms of mathematical formulae, exact or nearly so, within the limits of human observation."

To the present writer, the natural laws are convenient, useful statements of the observed correlations of phenomena. If a certain thing happens, another of a certain kind always has been observed to follow. This forms a "natural law," but nothing more than perfect correlation is meant or proved. Pragmatically it does not matter and can not be determined whether or not some or all of the natural laws are enforced by an inherent necessity. To deny a difference between well-grounded perfect correlation and a natural law is certainly not to question the utility of the natural law.

The combination of knowledge of correlations with knowledge of causal relations means, to me, merely a combination of knowledge of correlations with knowledge of other correlations. When the true nature of causation is grasped it can not mean more than this.

My third point, that the chain of causes must be broken off at some comparatively near finite point, is met by WRIGHT's just claim that in the method of path coefficients we are dealing with only that small part of the universe shown in our diagram of the lines of cause and effect. If more lines of cause and effect are taken into account the values of the path coefficients will be changed. A correlation coefficient is true no matter how many variables are considered. A path coefficient seems to be somewhat similar to a partial correlation coefficient with no plain indication of what variables are being held constant. In what way

could the method of path coefficients tell us more than we can easily find by the use of multiple correlation?

The statement (WRIGHT 1923, page 251) that I am attacking "the application of ordinary algebraic methods as by JENNINGS, FISH, PEARL, WENTWORTH and REMICK, and others" is ridiculous. I am defending ordinary methods against the method of path coefficients which I believe to be fundamentally unsound although at times it gives results in agreement with those ordinarily obtained.

Before closing, the writer wishes to make a point which is not essential to his argument but which indicates such a flaw in the method of path coefficients that it should not be omitted. In his reply to my original paper, WRIGHT (1923) gives an illustration of path coefficients which brings out the fact that a path coefficient may be greater than unity. This, which occurs in WRIGHT's figure 5, brings up the interesting question of what the coefficients of determination in this set-up will be. On page 245 WRIGHT states:

> "If one variable is completely determined by a number of others, the sum of the squares of the path coefficients leading to it, plus certain terms expressing joint determination by correlated variables, equals unity. A joint term of this kind is twice the product of the two path coefficients times the coefficient of correlation between the two variables in question. We have then:
>
> (2) $a^2+b^2+c^2+2bc\, r_{BC}=1$
>
> Because of this property the squares of the path coefficients give a useful measure of the degree of determination. Each one measures the portion of the squared standard deviation for which the factor in question is responsible."

When we have, as WRIGHT has, a path coefficient between X and $A=\sqrt{2}$ what may we say in regard to the degree of determination of A by X? Is it not $(\sqrt{2})^2$ or 2 on a scale which runs only from zero to unity? Or must a special rule be made in this case like the one which says that in deriving our equations we may go along the lines of causes directly from one variable to another, or backward and then forward, but never forward and then backward?

CONCLUSION

Until the objections urged in this and in my original paper (NILES 1922) are overthrown, the writer can not believe that the method of path coefficients is of the least value. It seems to be based upon a complete misapprehension of the nature of causation in the scientific sense.

LITERATURE CITED

NILES, H. E., 1922 Correlation, causation and Wright's theory of "path coefficients." Genetics 7: 258-273.

WRIGHT, SEWALL, 1923 The theory of path coefficients. A reply to Niles's criticism. Genetics 8: 239-255.

THE RELATION OF THE SPELT FACTOR IN WHEAT TO RACHIS INTERNODE CHARACTERS[1]

SARKIS BOSHNAKIAN

New York State Institute of Applied Agriculture, Farmingdale, Long Island

Received December 22, 1922

TABLE OF CONTENTS

INTRODUCTION 261
The characters and their determination 262
Density 262
Squareheadedness 263
Spelta and *sativum* characters 263
Description of parental material 265
Determination of genotypic make-up of the F_2 plants 265
The relation of spelting to density and squareheadedness 271
SUMMARY 274
LITERATURE CITED 275

INTRODUCTION

Triticum Spelta is one of the well-known and well-defined wheat species. In a cross between a spelt and a common wheat the usual segregation of characters takes place in the second and subsequent generations, but there are a number of characters peculiar to the species *Spelta*, which segregate out as a group, and not independently, as is the more common expectation. This grouping enables one to determine the numerical relationship of the segregates of the two specific classes, namely those of the species *Spelta* and of the species *sativum*. As will be inferred, we have here a case of specific segregation.

Of the specific characters the most striking is the brittleness of the rachis, which does not segregate out or recombine independently, but is very highly correlated with the glume form of the species *Spelta*. In the course of studies on the specific inheritance in wheat the author has examined several thousand F_2 and F_3 individuals issuing from various spelt × *sativum* crosses; he does not recall a single case where a plant had the *sativum* glume form and was at the same time brittle to any extent. It

[1] Paper 104, Department of Plant Breeding, CORNELL UNIVERSITY, Ithaca, New York. This study was made in coöperation with the Office of Cereal Investigations, UNITED STATES DEPARTMENT OF AGRICULTURE.

appears that the genes responsible for the spelt characters are located in the same chromosome, and that the loci of such factors as brittleness and glume form, just referred to, are relatively very close to each other.

Besides characters such as those mentioned above which may be regarded as linked, there are others which, as determined by proper genetic analyses, are located in other chromosomes, but are, nevertheless, affected by the factor in the presence of which the spike or head of the plant assumes the characteristics of the spelt type. Density and squareheadedness are internode characters which are independently inherited but are affected by the presence of the spelt factor. It is proposed in this paper to show the extent to which this spelt factor interferes with the full expression of the two above-mentioned rachis-internode characters.

The object of this paper is not to discuss the inheritance of each of these characters; although incidentally it has been necessary to consider their inheritance in the determination of the genotypic constitution of each of the F_2 plants used in these studies.

A full discussion of the inheritance of the spelt character will be found in a paper by LEIGHTY and BOSHNAKIAN (1921) and of density and squareheadedness in a recent publication by BOSHNAKIAN (1922). Except some preliminary studies reported in the latter publication, no paper has appeared as yet dealing with the correlation of the spelt character with density and squareheadedness, which is the theme of this report.

It is with pleasure that the author acknowledges his indebtedness to Dr. C. E. LEIGHTY of the Office of Cereal Investigations, UNITED STATES DEPARTMENT OF AGRICULTURE, as the material on which these studies were made was furnished through his courtesy. During the progress of this study the writer had also the privilege of sharing the valuable suggestions and criticisms of Professor H. H. LOVE of CORNELL UNIVERSITY, to whom he wishes to express his gratitude.

THE CHARACTERS AND THEIR DETERMINATION

Density

The wheat spike may be either compact or lax; dense heads are, as a rule, short; lax heads are long; intermediate forms are also to be found, but they are not as common as the extreme forms. In these studies density or compactness is expressed in terms of *average internode length*, which is found by dividing the length of the rachis (in millimeters) by the number of rachis internodes. Heads having an average internode length of about

2.0 mm or less are considered as dense; and those of about 3.0 mm or more are regarded as lax forms. These two terms, dense and lax, are, of course, relatively used.

Squareheadedness

This is a collective term used here to denote the presence of a number of spike characters which are caused directly or indirectly by the gradual shortening of the terminal rachis internodes. This character is expressed by the *coefficient of squareheadedness* which is found by dividing the number of internodes in the terminal third of the length of the rachis, by the number of internodes in the middle third of the rachis. Usually a head that may be regarded as a squarehead has a coefficient of squareheadedness of over 1.33. Any coefficient below this figure may be regarded as representing a non-squarehead form. The uses of these coefficients are described elsewhere by the writer (BOSHNAKIAN 1917).

Spelta and sativum characters

Most of the morphological characters of the spelt are contrasted by those of the *sativum* group; these characteristics, so far as they concern these studies, may be compared here briefly.

Spelts (figure 1 a) which do not carry the factor for density[2] are generally more lax than the lax *sativums*; the glumes are thick and stiff in texture, and strongly keeled to the base; the grains, being tightly held within the glumes, do not shatter; the rachis is very brittle.

The characteristics of the species *sativum*[3] (figure 1 b, c, d), on the other hand, are entirely the opposite. Its glumes are rather thin and soft; they are very weakly keeled; this keel does not run along the entire length of the glume; the grains are very loosely held and may be easily threshed out; the rachis is always tenacious. There are also other less striking contrasting characters between these two species, but they do not concern us at present.

Plants issuing from a cross between a spelt and a *sativum* are classified here with regard to the degree of spelting into classes 1, 2, 3, and so on, to 10, the classification depending upon the intensity of spelting of the heads as determined from their phenotypic appearance. A form intensely spelted is graded 1; those lacking the spelt characteristics altogether, namely, the *sativums*, are graded 10. Intermediate classes represent

[2] Commercial varieties of spelts are all lax. As a rule they do not carry the density factor.

[3] This specific name as used in this paper refers only to the common wheat types classified as *T. vulgare* Vill., *capitatum* Schulz., and *compactum* Host.

different grades of spelting. Class-10 individuals belong to the species *sativum* which comprises the following forms:

Vulgare (common wheat) average internode length usually more than 2.5 mm and squareheadedness less than about 1.33 (figure 1 c).

Capitatum (common squarehead) average internode length more than about 2.5 mm, squareheadedness more than about 1.33 (figure 1 d).

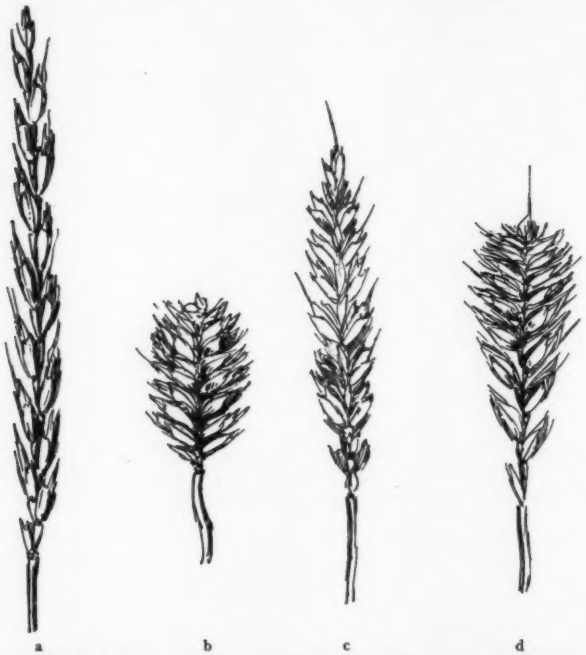


FIGURE 1.—The principal types of wheat referred to in the text. a, White Spelt; b, Dale Gloria; c, lax, non-squarehead *sativum* (*vulgare*); d, squarehead *sativum* (*capitatum*). The coefficient of squareheadedness of the latter is Sq=2.00.

Compactum (club wheat) average internode length less than about 2.5 mm. It may or may not be squareheaded (figure 1 b).

These three forms intergrade into each other, but they all have the spikelet form and, incidentally, other characteristics peculiar to the species *sativum*. The spelts may show to a degree some of their characteristics as regards density and squareheadedness, as we shall see, but as long as they exhibit the spikelet form of the species *Spelta* together with other incidental characters, they are regarded here as spelts.

Before taking up the question of the interrelation of these three characters, namely, spelting, density, and squareheadedness, it may be well to mention some conclusions which were arrived at in previous studies on similar material and which apply to the material to be considered. These conclusions are:

1. Density is partially dominant over laxness of the spike. It is caused by one factor difference.
2. Spelting is caused by one factor difference. This is partially dominant over the *sativum* form.
3. Squareheadedness is due to the interrelation of a number of growth factors capable of increasing and decreasing the intensity of the appearance of this character.

DESCRIPTION OF PARENTAL MATERIAL

The material studied (series 13255 a) represented the progeny of a cross between White Spelt (figure 1 a) and Dale Gloria (figure 1 b).

The internode characters of the parents were as follows:[4]

White Spelt *SSccq* D = 5.34 mm Sq = .86
Dale Gloria *ssCCQ* D = 1.41 mm Sq = 1.28

The spelt parent was lax; it lacked the factor for compactness, and carried none of the factors for squareheadedness, for, on other occasions when tested with non-squareheaded common wheats (*vulgare*), no squareheads were observed in their progeny. The non-spelt parent was a dense squareheaded club-wheat (*compacto-capitatum*) which carried the factor for compactness, and which upon being tested with non-squareheaded common wheats had shown that it carried factors capable of producing squareheadedness. This club, of course, could not carry the factor which produces spelting.

DETERMINATION OF THE GENOTYPIC MAKE-UP OF THE F_2 PLANTS

The heads of the first filial generation of this cross varied in degree of spelting from 3 to 5. The means of their internode length characters were D = 2.98 and Sq = 1.21. The parental and F_1 distributions with respect to these internode characters are shown in figure 2. The parental forms recorded are few but they give a general idea of their distributions with respect to these two characters.

[4] Average internode length or density is represented by D, and the coefficient of squareheadedness by Sq. The factors for spelting, compactness, and squareheadedness are *S*, *C*, and *Q*, respectively. The latter stands for a group of factors which are responsible for the production of squareheadedness.

Inasmuch as squareheadedness does not segregate in a definite ratio, for the present the inheritance of this material will be considered with respect to spelting and density alone.

With both *S* and C factors segregating in the 3:1 ratio, the F_2 population will segregate into the usual 9:3:3:1 ratio. The F_2 genotypes will behave in the F_3 in the following manner:

9 dense spelts
- 1 *SSCC* will yield all dense spelts
- 2 *SSCc* will yield 3 dense spelts : 1 lax spelt
- 2 *SsCC* will yield all 3 dense spelts : 1 dense *sativum*
- 4 *SsCc* will yield 9 dense spelts : 3 lax spelts : 3 dense *sativums*: 1 lax *sativum*

3 lax spelts
- 1 *SScc* will yield lax spelts only
- 2 *Sscc* will yield 3 lax spelts : 1 lax *sativum*

3 dense *sativums*
- 1 *ssCC* will yield dense *sativums* only
- 2 *ssCc* will yield 3 dense *sativums* : 1 lax *sativum*

1 lax *sativum*
- 1 *sscc* will yield lax *sativums* only

TABLE 1

Series 13255a White Spelt × Dale Gloria. Proportion of types obtained in the F_2 generation.

CLASSES	OBTAINED	CALCULATED	DEV.	P.E.	$\frac{\text{DEV.}}{\text{P.E.}}$
Dense spelts (*SC*).....	35	39.4	4.4	2.79	1.57
Lax spelts (*Sc*).....	14	13.1	.9	2.20	.41
Dense *sativums* (*sC*)....	15	13.1	1.9	2.20	.86
Lax *sativums* (*sc*).....	6	4.4	1.6	1.36	1.17

P = .667

The types obtained in the F_2 generation are shown in table 1. There were 49 spelts and 21 *sativums;* of the former 35 were dense and 14 lax; of the latter 15 dense and 6 lax. These results agree closely with the theoretical, as shown by the goodness of fit, where P = .667.

About half of the F_2 individuals were carried through the F_3. The proportions of genotypes in the F_2 and the types obtained in the F_3 are shown in table 2. In this table the data have been condensed from detailed measurements taken on each individual F_3 plant. It is intended to show the genotypic make-up of the F_2 plants rather than those of the F_3 themselves.

TABLE 2

Behavior of F_2 plants in the F_3. Series 13255a Turkey × Dale Gloria.

PEDIGREE NUMBERS	F_3 GENOTYPES AND THEIR PERFORMANCE	F_3 GENERATION DATA			F_3 GENERATION DATA CONDENSED								
					Sativums				Spelts				
		Spelting	Density	Square-headedness	Dense and semi-dense	Lax square-heads	Lax semi-squareheads	Lax *vulgare*	Very dense	Dense	Semi-dense	Lax	Very lax
	Dense spelts breeding true												
	SSCC												
2		2	3.11	.82	...	...	...	...	...	13	3	...	...
19		3	2.80	.86	...	...	...	...	...	25	...	...	...
	Dense spelts producing lax and dense spelt forms												
	SSCc												
4		1	3.67	.84	...	...	...	...	...	...	8	5	8
6		2	3.87	.92	...	...	...	...	...	12	10	6	11
	Heterozygous spelts producing only dense forms												
	SsCC												
17		7	1.79	1.31	3	...	...	...	6	2	2	...	...
24		2	2.32	1.01	10	...	...	...	8	8	9	...	...
25		9	2.35	1.45	10	...	...	...	24	9	...	1	...
26		7	2.45	1.18	12	...	...	...	18	6	1	...	...
30		7	1.81	1.30	3	...	...	...	21	8	2	...	...
	Heterozygous spelts producing dense and lax forms												
	SsCc												
1		2	2.33	1.13	2	1	...	...	...	7	3	2	...
7		3	2.93	1.00	3	...	...	...	4	13	2	7	9
10		2	2.77	1.13	4	...	1	...	3	2	5	4	6
15		5	3.11	1.01	14	1	...	1	6	9	10	8	6
16		4	3.13	1.13	6	1	...	...	...	8	6	5	6
21		5	3.10	1.15	4	2	...	1	7	9	11	10	8
	Lax spelts breeding true												
	SScc												
14		1	4.89	.92	...	...	...	...	...	...	...	...	16
3		1	4.00	.84	...	...	...	...	...	...	...	...	6

TABLE 2 (continued)

PEDIGREE NUMBERS	F_3 GENOTYPES AND THEIR PERFORMANCE	F_3 GENERATION DATA			F_3 GENERATION DATA CONDENSED								
					Sativuma				Spelts				
		Spelting	Density	Square-headedness	Dense and semi-dense	Lax square-heads	Lax semi-squareheads	Lax *vulgare*	Very dense	Dense	Semi-dense	Lax	Very lax
	Heterozygous spelts producing only lax forms *Sscc*												
8		3	4.40	1.00	...	2	1	1	...	...	...	8	7
11		1	4.76	.91	...	1	7	...	...	...	...	...	25
23		7	4.89	1.00	...	16	...	...	...	...	...	...	34
28		4	4.68	1.13	...	6	...	11	...	...	1	4	30
	Dense *sativums* (*T. compactum*) breeding true *ssCC*												
20		10	1.63	1.03	20	...	...	...	...	...	...	...	...
5		10	1.86	1.00	10	...	1*	...	...	...	...	...	...
	Heterozygous dense *sativums* *ssCc*												
9		10	2.77	1.13	20	2	3	1	...	...	...	...	...
12		10	1.72	1.37	10	...	3	...	...	...	...	...	...
13		10	1.86	1.43	9	4	1	2	...	...	...	...	...
32		10	2.22	1.55	23	...	6	4	...	...	...	...	...
	Lax squareheads producing squareheads and non-squareheads *ssccQ*												
22		10	4.00	1.67	...	31	...	7	...	...	...	...	...
29		10	4.17	1.53	...	10	13	7	...	...	...	...	...
27		10	3.37	1.58	...	8	6	9	...	...	...	...	...
31		10	3.15	1.62	...	16	3	1	...	...	...	...	...
18		10	3.47	1.76	...	6	5	10	...	...	...	...	...
	Lax *vulgare* (non-squarehead) breeding true *ssccq*												
33		10	4.40	1.15	...	...	...	14	...	...	...	...	...

*Undeveloped.

Under the general division "*sativums*," four subdivisions are made to give an approximate idea of the types obtained with respect to their internode characters. The first column under *sativums* contains all the dense and semi-dense *sativums;* this is a genetic classification and includes all the *ssCC* and *ssCc* types; in very rare instances it may contain a few forms of the *SsCC* type which represent spelts of such dilute form (between classes 9 and 10) that they cannot be distinguished from the dense *sativum* forms. The chances of this wrong classification being somewhat remote, this source of error may be disregarded for the present at least. The second, third and fourth columns under "*sativums*" represent phenotypic and arbitrary classifications of the lax forms; they represent degrees of squareheadedness approximately of 1.40 or more, 1.20 to 1.40, and less than 1.20, respectively.

The subdivisions under "spelts" contain five phenotypic classifications with respect to density alone. The plants classified under the first and second columns represent all the *SSCC* and *SsCC* individuals and the great majority of *SSCc* and *SsCc* genotypes. The "very dense" class is usually made up of spelts which were graded 8, 9 and a few which were intermediate between classes 9 and 10. In this last case the degree of spelting is so dilute that the plants resemble somewhat the *sativum*, and like the latter show more or less the total effect of the *C* factor. So far as the presence of the *C* factor is concerned the plants in the "very dense" class do not differ genotypically from the plants in the next column which contains the "dense spelts." The subdivision "lax" is likely to contain an appreciable number of *SsCc* forms. The forms in the "very lax" class may be considered to be free of the *C* factor.

All the 33 plants, except the lax *sativums* tested in the F_3, were selected at random. Care was taken, however, to test all the 6 *sativums* to note their behavior with regard to density and squareheadedness. The 33 plants tested fell into the nine expected theoretical classes; the numbers of F_2 plants belonging to these classes as shown in table 2, are condensed in table 3. In five of these classes, the numbers obtained were as close to the theoretical as could possibly be expected; in two classes (*SSCc* and *SsCc*) there was a theoretical deficiency of 2.13 and 2.58 individuals, respectively.

In the case of the lax *sativums* six plants were obtained where but two were expected. Of these six plants one was a stabilized extracted *vulgare* (*ssccq*). In calculating the expectations for *vulgare* it was supposed that this form would arise in one case out of four, but probabilities are that its occurrence is more rare. The goodness of fit of the data as they stand is

only P = .334. The failure of agreement of the experimental with the theoretical numbers is due to the excess of *sativums*. It was mentioned that while all the plants were selected at random, care was taken to grow all the lax *sativums*. Had there been no personal factor introduced in the selection of all the F_2 *vulgare* plants the random sample should theoretically contain lax *sativums* in the proportion in which they were found in

TABLE 3

Series 132552. Condensed table showing the number of the F_2 plants which proved to belong to the theoretical classes. This table is based upon the F_3 data presented in table 2.

GENOTYPES	OBTAINED (UNCORRECTED)	CALCULATED	DEV	OBTAINED (CORRECTED)	CALCULATED	DEV.
SSCC	2	2.06	.06	2	1.87	.13
SSCc	2	4.13	2.13	2	3.75	1.75
SScc	2	2.06	.06	2	1.87	.13
SsCC	5	4.13	.87	5	3.75	1.25
SsCc	6	8.58	2.58	6	7.49	1.49
Sscc	4	4.13	.13	4	3.75	.25
ssCC	2	2.06	.06	2	1.87	.13
ssCc	4	4.13	.13	4	3.75	.25
ssccQ (Squareheads)	5	1.55	3.45			
ssccq (*vulgare*)	1	.51	.49			
sscc (in F_2 proportion)				2.82	1.87	.95
	P = .334			P = .989		

the F_2, or 2.8 plants instead of six. Making this correction for the lax *sativums* the calculated value of the goodness of fit would be P = .989, which shows a very close agreement between the expectations and the experimental results.

The proportions of dense and lax F_2 plants of the spelt and *sativum* forms which showed segregation with respect to these characters are presented in table 4, which in a general way bear out the hypothesis. When the segregation was between spelt and *sativum* without regard to density, or between dense and lax forms within the *sativum* group, the agreement between experimental results and expectations were satisfactorily close.

In the dense and lax classes within the spelts, however, marked discrepancies are found. There is always a very significant excess of lax spelts. It is very likely that these "lax" spelts, which were distinguished from the "very lax" class, contain an appreciable proportion of spelts heterozygous for the density factor. Reference to this possibility has

been already made. The density frequency curves in spelts, unlike those among the *sativums*, are more or less continuous and appear more so when the population is small, so that it is not possible to separate the lax

TABLE 4

Behavior of F_2 plants heterozygous for C, or S, or for both factors. Series 13255a.

GENOTYPES OF THE F_2 GENERATION	NUMBER OF HETEROZYGOUS PLANTS TESTED	TYPES OF PLANTS OF THE F_3 GENERATION							
		Spelts				*Sativums*			
		Dense		Lax		Dense		Lax	
		Obtained	Calculated	Obtained	Calculated	Obtained	Calculated	Obtained	Calculated
SSCc	60	30	45.0	30	15.0	....	....	....	....
SsCC	163	125	122.3	....	....	38	40.7	....	....
SsCc	207	100	116.4	66	38.8	33	38.8	8	12.9
Sscc	154	....	....	109	115.6	....	....	45	38.5
ssCc	88	....	....	....	....	62	66.0	26	22.0
Total	672								

spelts from those which are heterozygous for the *C* factor, especially those forms which are intensely spelted, as these forms practically neutralize the effect of a single dose of the *C* factor.

THE RELATION OF SPELTING TO DENSITY AND SQUAREHEADEDNESS

Having analyzed and determined the genotypic make-up of each of the individuals of the F_2 generation, we can consider now the distribution of the correlation scheme in figure 2 to determine the relationship of the spelt character to density and squareheadedness.

This table represents in a diagrammatic way the quantitative and qualitative distributions of the individuals with respect to the three characters under consideration. The squares represent classes of a correlation table, one magnitude representing degrees of density, as determined by the average internode length; and the other, of squareheadedness. Each small circle stands for a plant and the different degrees of shading represent different degrees of spelting as indicated in the legend. The numbers 1 to 33 stand each for the pedigree number of a plant which was tested in the F_3. To know the exact F_2 and F_3 behavior of these plants, the reader can refer to table 2.

By taking into account the genetic behavior of each F_2 plant tested and locating it on this correlation table (figure 2) it was noted that like

genotypes with respect to the *S* and *C* factors were grouped within more or less well-defined areas in this table. With the aid of the F_3 data in table 2, these areas have been mapped out and the genotypic formula of each is shown.

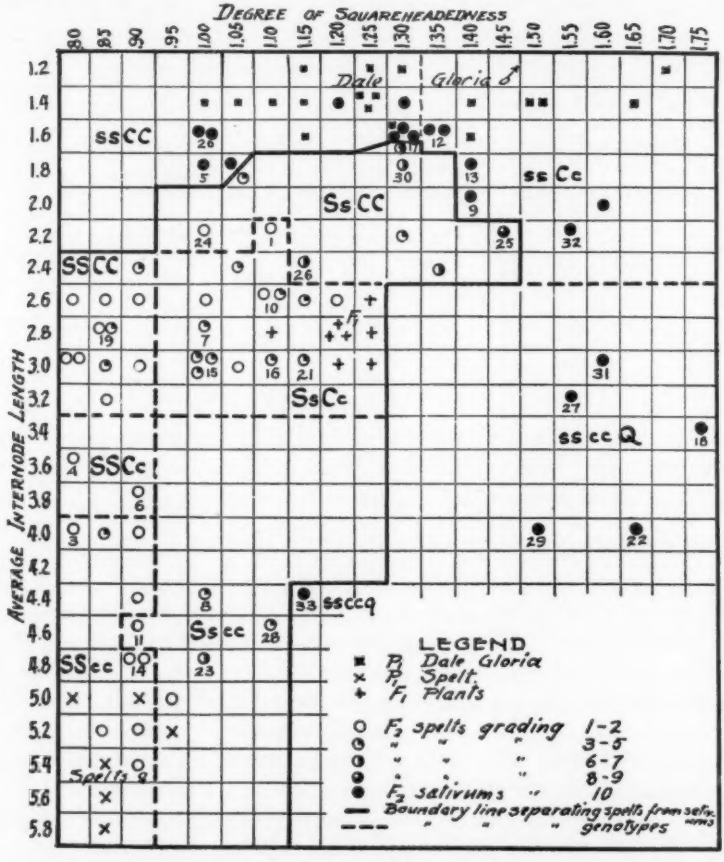


FIGURE 2.—Correlation table showing the isolation of spelt segregates from *salivums* of the same generation, and also the peculiar groupings of like genotypes.

As mentioned, these boundaries were determined from the F_3 data of the plants tested; those which were not tested and happened to fall in one of these areas are not necessarily of the same genetic constitution as those which were tested.

Let it be recalled that the distributions of the F_2 plants in this table were made according to density and squareheadedness alone. If there were no correlation between the *S* factor and the above two internode characters, both spelts and *sativums* would be expected to have the same type of distribution over this area. The presence of a number of definite groupings, however, shows that there are certain relations between the spelt character, and density and squareheadedness.

The following may be regarded as the chief characteristics of this table:

1. The plants carrying the *S* factor occupy that portion of the table which contains forms that are neither highly dense nor squareheaded. All the *sativums* are grouped outside this area.

2. Regarding the relation of spelting to squareheadedness it may be noted that within the area occupied by the spelts, the homozygous *SS* forms occupy a section at the extreme left-hand side of the table where the squareheadedness coefficient is less than .90, whereas the coefficient of the heterozygous spelts is always above .90.

3. As to their arrangement with respect to density, the density of the spelts is, generally speaking, proportional to the dose of the *C* factor they carry. But density is not merely dependent upon the *C* factor but also upon the *S* factor. To illustrate this we may take the groups where the *C* factor is constant, in one case in a homozygous form, and in the other case in a heterozygous condition. The ranges of density of the different tested F_2 forms shown in the table are as follows:

ssCC	1.4 to 1.8 mm	*ssCc*	1.6 to 2.2 mm
SsCC	1.6 to 2.4 mm	*SsCc*	2.2 to 3.0 mm
SSCC	2.4 to 3.0 mm	*SSCc*	3.4 to 3.8 mm

It is seen above that the *C* factor remaining constant, the greater the dose of the *S* factor the laxer the form. The *S* factor then acts as a partial inhibitor of density. This *S* factor may also be regarded as a factor increasing the rachis internode length. In the instance under consideration both views may be regarded as correct.

That the *S* factor acts as an inhibitor may be demonstrated by crossing the variety Black Bearded Spelt with a lax *sativum*. The F_2 population will contain a large number of clubs, not all as dense as Dale Gloria, yet dense enough to be regarded as clubs. Of these dense forms none will be a spelt. In this case the density factor was introduced by the lax spelt parent.

To show that the *S* factor acts as a lengthening factor, a lax *sativum* may be crossed with a lax spelt (similar to the White Spelt used as the

spelt parent of the cross discussed in this paper) in which none of the F_2 individuals will be dense. The F_2 population will consist of lax spelts and lax *sativums* (*vulgare*). The spelts, however, will be considerably more lax than the *vulgare* forms. This phenomenon may be noted also when the densities of the *SScc* and *sscc* forms (figure 2) are compared. The lax spelts are considerably laxer than the lax *sativums*.

The study of the characteristics of the distributions of *sativums* and spelt populations presented in this correlation table leads to the conclusion that the factor for spelting acts as a partial inhibitor for both density and squareheadedness.

SUMMARY

The purpose of this paper is to show in a quantitative way the relation of the spelt character to density and squareheadedness.

The material in this analysis consists of an F_2 population of a cross between a dense *sativum* (club) and an ordinary (lax) spelt. These F_2 plants have been tested in the F_3 and their genotypic make-up determined.

While it is not intended to present here the inheritance of each of the characters concerned, namely, spelting, density, and squareheadedness, it has been found necessary to refer to their mode of inheritance in order to determine the genotypic make-up of the plants to be used in the correlation studies.

So far as this material is concerned, the factors for density and spelting are found to be partially dominant over those producing laxness and *sativum* characters, respectively. The F_2 segregation curve of density being bimodal and that of spelting, of the U-type, the curve of the heterozygous populations merging in each case into that of the homozygous dominant individuals. As to squareheadedness, it appears to be due to more than one growth factor, yielding, therefore, no definite segregation curve.

Using a correlation table in which one magnitude represents different classes of density, and the other magnitude, various classes of squareheadedness, studies are made of the distributions on this correlation area of plants whose genotypic constitution with respect to the spelt (*S*) and density (*C*) factors are definitely known.

It is noted that plants of like genotype with respect to the density and spelt factors congregate within rather definite areas on this correlation table; the reasons for this type of segregation as shown in the text are the following:

(a) Within both *sativum* and spelt classes the degree of density depends upon the presence and dose of the *C* factor.

(b) The *C* and squareheadedness factors remaining constant, whether in homozygous or heterozygous condition, plants which lack the *S* factor (*sativums*) are decidedly more dense and more squareheaded than those carrying the *S* factor (i.e., spelts).

(c) The *C* and squareheadedness factors remaining constant, plants carrying a single dose of *S* are more dense and more squareheaded than plants which carry a double dose of the *S* factor.

It is concluded that the spelt factor decidedly interferes with the full expression of the factors for density and those for squareheadedness.

LITERATURE CITED

BOSHNAKIAN, S., 1917 The comparative efficiency of indexes of density and a new method for determining squareheadedness in wheat. Jour. Amer. Soc. Agron. **9**: 231–247.

1922 The genetics of squareheadedness and of density in wheat, and the relation of these to other characters. Cornell Univ. Agric. Exp. Sta. Memoirs **53**: 801–882.

LEIGHTY, CLYDE E., and BOSHNAKIAN, SARKIS, 1921 Genetic behavior of the spelt form in crosses between *Triticum Spelta* and *Triticum sativum*. Jour. Agric. Res. **22**: 335–364.

THE GENETICS OF CURLY WING IN DROSOPHILA. ANOTHER CASE OF BALANCED LETHAL FACTORS[1]

LENORE WARD

4404 Morningside Avenue, Sioux City, Iowa

Received January 3, 1923

TABLE OF CONTENTS

INTRODUCTION 276
Acknowledgments 277
Source of curly 277
Description of curly 278
Dominance, heterozygosity, and the production of twin hybrids 280
Chromosome involved 283
Balanced lethal nature of curly stock 284
Crossover factors in the second chromosome with the gene for curly 286
Effect of the crossover factors on chromosomes I and III 289
Increase of crossing over due to a higher temperature 293
Crossover factors in the homologue of the curly chromosome 295
Effects of temperature on the crossover factors in the homologue of the curly chromosome 295
Locus of the gene for curly 296
Cinnabar[2], a cinnabar allelomorph in the curly chromosome 297
Evaluation of curly 298
SUMMARY 299
LITERATURE CITED 300

INTRODUCTION

Curly wing, with which this paper deals, is the first easily and sharply separable dominant wing character that has been found in the second linkage group of *Drosophila melanogaster*. It maintains itself in a balanced lethal stock that has arisen during selection. In out-crosses this stock produces twin hybrids which offer a parallel to cases in Oenothera; and the inbred stock occasionally yields, on account of crossing over, apparent vestigial and cinnabar[2] mutants. Analysis of this stock shows that one second-chromosome homologue carries the gene for the dominant character curly. In homozygous condition the curly gene acts as a lethal, except in rare cases in individuals which survive as dwarfed but fertile flies that are very curly.

[1] Contribution from the Zoölogical Laboratory of the UNIVERSITY OF MICHIGAN.

The homologue that carries curly carries also a new recessive eye color, namely, cinnabar[2], which is a less extreme allelomorph of the original cinnabar. The other homologue carries a recessive character, vestigial, and, about ten units to the left of vestigial, a recessive lethal. The vestigial flies are nearly all sterile, which may mean that the vestigial is a new nearly sterile allelomorph of vestigial (viz., v_9^2), or that there is a sterility factor also in the homologue with vestigial.

In each of the two second chromosomes, there are dominant factors which inhibit crossing over along the entire length of the second chromosome under normal temperature conditions. Very rarely a case of crossing over does occur at normal temperatures; but at high temperatures ($30° \pm$) crossing over in the right half of chromosome II becomes fairly frequent (6 percent $\pm$). Double and triple crossovers form a disproportionately large percentage of the total of such crossovers. Curly, or more probably the crossover suppressors that are in the chromosome with curly, appear to reduce the amount of crossing over in chromosomes I and III. It is probable that the crossover modifiers in the vestigial homologue are identical with those in the curly homologue, and that curly arose in a stock homozygous for these modifiers.

The excellent viability of curly, its ease of identification, dominance, and property of preventing crossing over in the second chromosome at normal temperatures make it a useful character in experiments and in the "balancing" of other stocks that would otherwise require selection each generation.

ACKNOWLEDGMENTS

This investigation was carried on under the direction of Professor A. F. SHULL, and I am indebted to him for much assistance. Dr. H. J. MULLER has rendered valuable aid throughout the work, and Professor ROY E. CLAUSEN has made many valuable suggestions. The paper has also been examined by Dr. C. B. BRIDGES, who has checked up the linkage data, and has suggested new interpretations for some of the findings.

SOURCE OF CURLY

The original wild stock from which curly came was collected in Ann Arbor, Michigan, in September, 1918, and kept for use as wild stock. It eventually produced three entirely wingless females. These females were crossed to wild and the F_1 were all normal flies. The inbred F_1 produced an F_2 chiefly wild, but a few with crumpled wings. These abnormal-winged flies were mated together, and again gave wild and abnormal flies.

This practice of selecting the abnormal-winged type of fly each generation, for matings, was continued, with increasing proportion of abnormal wings of several different types, until abnormal types constituted about 50 percent of the total number of flies from each mating. There were wings resembling club, jaunty, cut, truncate, apterous, beaded, balloon, vestigial, and blistered, and some flies with no wings at all. None of these forms except vestigial could have been due to contamination from other mutant-wing stocks because up to that time none of these other mutants had been brought to the laboratory at the UNIVERSITY OF MICHIGAN.

A definite selection experiment was carried on from March to November, 1920, in an attempt to establish a race of some one definite type of wing character from this mixture, but this attempt was unsuccessful. Flies somatically identical, when bred together, failed to produce offspring that resembled the parents any more closely than did the general population of abnormal flies. Of a total of 6769 flies from matings of like abnormal parents, 5881 normal flies and 1588 abnormal flies were produced.

The curly-winged type of fly appeared among the irregular offspring; and because it seemed a more stable character than any of the others, it was isolated and inbred. At first normal flies appeared in the stock bottles with curly. However, after nearly a year of mass selection for the curly character with an ever-decreasing proportion of the normal-winged type, a 100-percent stock of curly flies was obtained. No normal flies appeared after that until July, 1921. Investigation of the apparently normal individuals which appeared at that time demonstrated the wild type to be only phenotypically wild, but genetically curly. Such flies are found only in very old dry culture bottles.

In the stock cultures of curly there appeared an occasional vestigial-like or cinnabar-like fly. This phenomenon of the occurrence of vestigial and of cinnabar (c_n^2) may, as is shown later, result from a rare crossing over between a lethal factor linked to vestigial in the homologue of the curly chromosome. Cinnabar[2] also appears on account of crossing over. It was conceivable that vestigial might have been introduced into the stock by contamination, since stocks of vestigial flies were at that time in use in the laboratory. However, this is not likely. The genetic analysis points rather to the occurrence, by mutation, of a new allelomorph of vestigial (v_g^2), more extreme than vestigial and nearly always sterile.

DESCRIPTION OF CURLY

The general size, shape, and venation of the wings of the average curly flies is essentially that of the wild, but the color of the wing membrane is

paler, and the texture is thinner in curly flies. The veins are much less rigid, and finer, and of a lighter color. There is a tendency toward a scallop on the inner margin of the wing at the termination of the fifth longitudinal vein, and there is usually one long fine wrinkle across the wing from near the termination of the second longitudinal vein to the intersection of the fifth longitudinal vein and the margin of the wing. Sometimes there are two shorter crossed wrinkles instead of the single long wrinkle, and not infrequently several short fine wrinkles extending between the third and fourth longitudinal veins (figures 1 and 2).

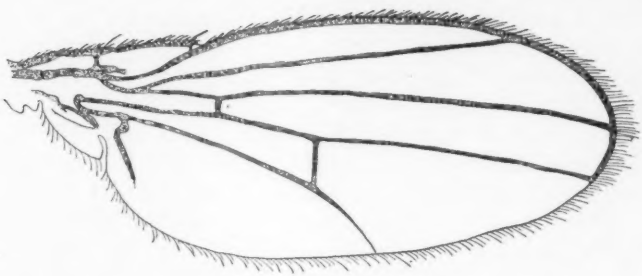


FIGURE 1.—Normal wing of *Drosophila melanogaster*.

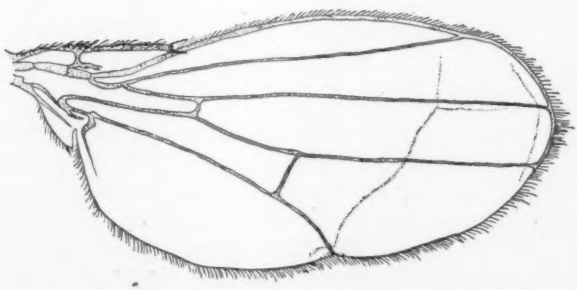


FIGURE 2.—Curly wing, showing cross and longitudinal wrinkles.

The degree of curvature of the wings of curly flies varies from a slight upward bend of the wings at the tip to an extreme roll of 360 degrees or over (figure 3). When the wings are very much curled there are more wrinkles in them, and these extra wrinkles extend both crosswise and lengthwise of the wings.

DOMINANCE AND HETEROZYGOSITY OF CURLY, AND THE PRODUCTION OF TWIN HYBRIDS

The first test matings using curly were crosses to wild. Either reciprocal cross yielded an F_1 invariably showing a 1 : 1 ratio of curly and

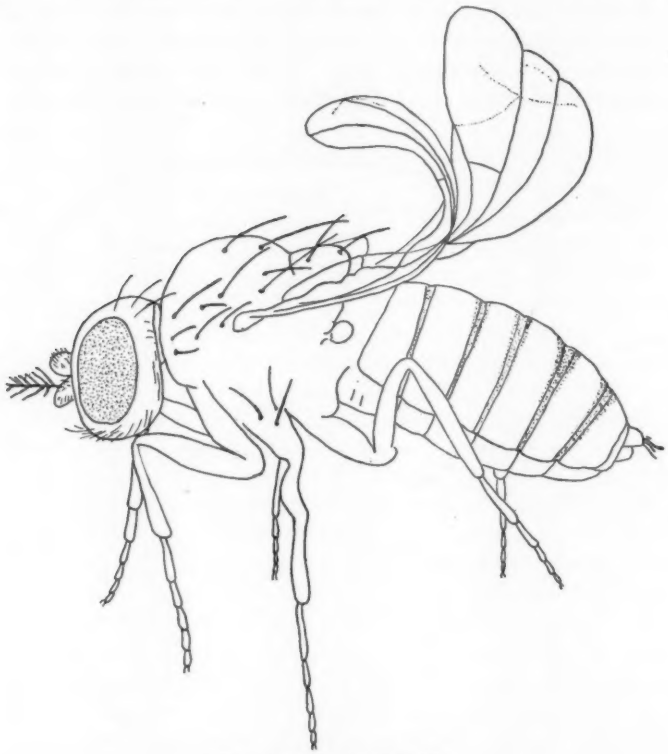


FIGURE 3.—Curly-winged fly.

wild, which is typical of a simple dominant heterozygous character (table 1).

TABLE 1

Offspring from cross of curly to wild.

PARENTS	F_1 CURLY	F_1 WILD	TOTAL
Curly × wild.....	1023	991	2014

Additional proof that curly is a heterozygous form was obtained from a mating of the dominant heterozygous star eye with curly. A curly female bred to a star-eyed male gave in F_1, curly, star, curly star, and wild, just as would be expected were curly a heterozygous dominant character. Further test matings of the curly star F_1 flies together gave an F_2 of curly star flies only. All of these results indicate that curly is a heterozygous form and is lethal when homozygous. BRIDGES informs me that he has raised cultures from F_1 curly stars in which C_y not-*S* flies occurred in considerable numbers. These C_y flies, which had c_a^2 eye color, were smaller than normal, had especially pinched abdomens, and especially curled wings. When they were out-crossed they were fertile, and all of the offspring were curly of the usual type. This shows that curly can occasionally survive in the homozygous form, and belongs to the class of "semi-lethals."

Curly flies were regarded as simple dominant heterozygotes until a cross was made between vestigial and curly. An F_1 from curly and vestigial gave, not the expected 1 : 1 ratio of curly and wild, typical of the hybrid ratio of a dominant heterozygote and a recessive, but a 1:1 ratio of curly and vestigial (table 2). This result appeared to indicate that the curly fly was heterozygous for vestigial as well as for curly.

TABLE 2

Offspring from the cross of curly and vestigial.

PARENTS	F_1 CURLY	F_1 VESTIGIAL	TOTAL
Curly × vestigial.....	512	508	1020

It is later shown that curly is a second-chromosome character, and the suggestion immediately presented itself that a vestigial gene existed in the chromosome homologous to that of curly. To test this hypothesis an F_1 wild male from a cross of curly by wild was mated to a vestigial female from COLUMBIA stock. This F_1 wild male must have contained in it the homologue of the curly chromosome. This cross produced vestigial and wild flies in equal numbers, hence a gene for vestigial must have been carried in the homologue of the curly chromosome.

If the curly stock is heterozygous for vestigial, we might expect to obtain in a culture bottle of curly stock about one-third of the flies with vestigial wings. As a matter of fact the curly stock maintains itself phenotypically pure for curly with no vestigial flies appearing except in rare cases.

These vestigial flies have proved to be sterile in nearly all cases. Of 24 matings (table 3) in which they were used, 6 were fertile and 18 were sterile, and the fertile ones, which were females, hatched only a few (84) offspring.

TABLE 3

Results of the 6 successful matings out of a total of 24, in which only one of the parents in each case was a vestigial segregant.

BOTTLE	PARENTS	OFFSPRING			
		Wild	Vestigial	Cinnabar	Curly
1	v_g ♀ × cn^1 ♂	3	0	0	0
9	v_g ♀ × cn^1 ♂	26	0	0	0
13	v_g ♀ × C_y ♂	0	10	0	16
16	v_g ♀ × C_y ♂	0	0	0	2
21	v_g ♀ × C_y ♂	0	0	0	11
24	Real wild ♀ × v_g ♂	16	0	0	0
Totals	84	45	10	0	29

No males were found to be fertile. Professor CLAUSEN reports that he has never found any of these vestigial segregants fertile. This sterility suggests that the mutant is not vestigial itself, but a new and sterile allelomorph (v_g^2).

The failure of the appearance of vestigial flies in the curly stock in large numbers might be explained as the result of the presence of a recessive lethal linked to the gene for vestigial in the homologue of the curly chromosome. Their occasional appearance in the curly stock could be explained by the assumption that they are the result of crossing over between the lethal and vestigial, which thus allows the homozygous vestigial flies to appear. Stock bottles kept at normal room temperature show fewer of these vestigial crossovers than do bottles kept at 30° centigrade in the incubator. As is demonstrated later, high temperature increases the amount of crossing over in the second chromosome, and this fact fits in well with the above hypothesis of the origin of the vestigial flies in the curly stock.

When curly from the balanced stock is crossed to wild, it produces in the F_1 twin hybrids of the types curly and wild. If these types are inbred

separately, the not-curly flies give only not-curlies in each successive generation, but the curly splits up into curly and wild. This parallels the case in Oenothera, where *Oenothera Hookeri* × *Oe. Lamarckiana* produces twins of the types *laeta* and *velutina*, the *laeta* from these crosses splitting in each successive generation into *laeta* and *velutina*, while *velutina* has a constant progeny (DE VRIES 1918).

CHROMOSOME INVOLVED

Since it is established that curly is heterozygous and that vestigial or a vestigial allelomorph is located in its opposite chromosome, and since vestigial is a second-chromosome character, it is evident that curly is in the second-chromosome linkage group. Positive experimental tests to disprove the linkage of curly to first- and third-chromosome genes and to prove the linkage of curly to second-chromosome genes were undertaken.

If the gene for curly were located in the first or X chromosome, then, because curly has been shown to be a dominant heterozygote, a mating between a curly male and a wild female would produce an F_1 of curly daughters and wild sons. That is not the case. In such a cross the sexes are found to be equally distributed between curly and wild flies. The same holds true in a cross of a curly male and an "Xple"[2] female; that is, a female which is homozygous for the first-chromosome characters, scute, echinus, cut, vermilion, garnet and forked. Here also the curly is found to be present in half of the F_1 males which all show the above six recessive first-chromosome characters. Curly therefore exhibits complete independence of the first or X chromosome (see table 7).

If the gene for curly were located in the third chromosome, it would be found to show linkage with other genes in that group. A back-cross test of curly and the homozygous recessive multiple stock called "IIIple," carrying the third-chromosome recessives, roughoid, hairy, scarlet, peach, spineless, and sooty, was made. As might be expected, the F_1 were one-half curly and one-half wild-type, both types heterozygous for the third-chromosome genes. F_1 curly females were mated to "IIIple" males and the back-cross generation showed complete independence of curly from each of the characters, roughoid, hairy, scarlet, peach, spineless and sooty (see table 8). Because these genes have loci distributed throughout the third chromosome, it is scarcely possible that any linkage with curly would have failed of detection.

[2] Since it is customary for workers in Drosophila to use the terms "Xple," "IIple," and "IIIple" to designate certain mutant stocks, such terms are used here.

An experimental test for the relation of curly to the second-chromosome group involved a septuple back-cross of curly with the homozygous recessive multiple second-chromosome stock called "IIple," which is homozygous for dumpy, black, purple, curved, plexus, and speck. Curly crossed to "IIple" gave an F_1 one-half of which were curly and one-half wild-type, both heterozygous for the "IIple" characters. F_1 curly females were then back-crossed to "IIple" males; and there were produced in the first back-cross generation only curly and "IIple" flies, in the ratio 1:1. This experiment involved many different attempts, and a great number of flies, and always gave the same result.

There was manifested a striking lack of crossing over, for among 2487 back-cross individuals (see table 4), not a single crossover fly appeared. The gene for curly, or some other gene linked to it, modified the crossing over of the "IIple" genes and kept them distinctly associated. The total lack of recombinations of curly with each of these second-chromosome characters showed that curly is a member of the second linkage-group. Crossover flies were finally obtained from curly and "IIple" by subjecting the heterozygotes to a high temperature, as described later in this paper.

BALANCED LETHAL NATURE OF CURLY STOCK

It has been shown that curly is ordinarily lethal when homozygous; likewise, the vestigial allelomorph is linked to a lethal. Since two lethals are present, one in each of the two homologous chromosomes, and since crossing over is practically non-existent in this stock at ordinary temperatures, a condition of "balanced lethals" obtains. This situation is similar to that described by MULLER (1919) in the case of beaded wing in Drosophila, and somewhat similar to the situation described by FROST (1920).

Dominant characters that are lethals when homozygous are more easily detected when they occur than recessive lethals. In nature, mutations which are lethal probably arise very frequently and then vanish on account of the process of natural selection. If we suppose, however, that lethals are not left to natural selection, but are arbitrarily selected for by the conditions of an experiment, we may expect them to survive, provided survival is possible.

If some dominant heterozygous character such as curly wing is being selected for, the chances of the survival of any lethal arising in its chromosome or in the one homologous to it are good because of the condition of the "enforced heterozygosity" of the stock. In the selection experiment described above, flies were chosen each generation for the curly wing character and thus kept in such an enforced heterozygotic condition.

Lethals would occasionally arise, and would persist because the stock was kept heterozygous. In fact, if the lethal curly character were selected for and another lethal had arisen in the chromosome homologous to it, that other mutant lethal would be more likely to be saved along with the curly lethal than would the normal gene, because stock containing it would throw a higher percentage of curly flies. In this manner there would be an automatic establishment of a "balancing lethal" for the curly lethal, as there is in selecting for any heterozygous stock in respect to one lethal. If flies containing both of these lethals happen to be taken for parents, a self-perpetuating stock of "balanced lethals" is the result and the stock maintains itself pure for its hybrid constitution.

If the population were only ordinarily hybrid and not balanced lethally, a separation into its respective homozygous strains might be expected to result when the stock is inbred. Flies of the "balanced lethal" type, however, remain necessarily heterozygous for the two lethals and for any other of the genes located in the same chromosome with those lethals unless the balancing effect of the lethals is disturbed by crossing over.

In rare cases mentioned above, in stocks of curly flies which have otherwise continued pure for curly for many generations, it has been noticed that a few vestigial or cinnabar^2 flies appear. This appearance of the vestigial might be explained if crossing over took place between a gene for vestigial, located in the homologue of the curly chromosome, and a lethal which is present in the same homologue with it. The supposition would be that in the Curly stock, once in several hundred times, crossing over does occur in the chromosome between the gene for vestigial and for a linked lethal. Having thus rid itself of the lethal by crossing over, the vestigial appears. The occurrence of the cinnabar^2 flies is made possible also by crossing over. It is shown later in this paper that a gene for cinnabar^2 is present in the chromosome with curly. When crossing over occurs between the gene for cinnabar^2 and curly, cinnabar^2 rids itself of one curly gene and can appear in homozygous form in a fly that carries both curly and cinnabar^2 in one homologue, and only cinnabar^2 in the other. Such flies, by inbreeding, would produce cinnabar^2 flies free from curly. Such cinnabar flies, free from curly, have been found and can not be explained as surviving homozygous curly cinnabar.

This phenomenon of the appearance of the vestigial and cinnabar^2 flies in the curly stock simulates mutation, although it is not really such. In the case of vestigial it is merely the effect of the combination of the two recessive vestigial genes, one of which has lost a linked lethal. In the case of the cinnabar^2 it is the effect of the combination of the two

recessive cinnabar[2] genes, one of which has crossed over from curly. The curly cinnabar[2] flies are in some cases crossovers ($C_y\ c_n^2$ in one II, c_n^2 in the other), and in other cases are surviving homozygotes ($C_y\ c_n^2$ in both homologues).

The same type of apparent mutation phenomena as the production of vestigial and cinnabar[2] flies in the "balanced lethal" curly stock had been shown to exist in other "balanced lethal" stocks. MULLER in 1918 demonstrated that the "balanced lethal" beaded stock in Drosophila contains recessive factors that sometimes cross over from their linked lethals and become visible because of their homozygosity. On the basis of the beaded results he proposed the same kind of explanation for mutation in Oenothera. FROST (1919) found that the same facts hold for Matthiola, except that one factor there is not really lethal. G. H. SHULL (1923) states that some of the Oenothera mutations also arise by the crossing over of genes linked to lethals which are in a balanced state.

MULLER'S theory for the Oenothera mutations is supported by evidence from the beaded case, in which the recessive factors which exhibited apparent mutation were artificially inserted. The evidence from curly lends even more support to his theory because in curly the recessive genes arose in the stock in the natural course of selective breeding, indicating the general nature of this phenomenon and its natural occurrence in balanced stocks. Mutation phenomena of the same kind are hence to be expected in nature and must be reckoned with.

CROSSOVER FACTORS IN THE SECOND CHROMOSOME WITH THE GENE FOR CURLY

Since the gene for the production of curly wing is in the second chromosome, the best cross to use for testing the linkage relations of the curly factor to the other factors in the second chromosome is one involving the homozygous recessive multiple second-chromosome stock called "IIple." This stock contains the factors for dumpy wing, located at 9 on the second chromosome, black body at 47, purple eye at 53, curved wing at 78, plexus wing venation at 98, and speck body at 105. These "IIple" genes give checks on six different loci of the second chromosome.

A curly female was first crossed to a "IIple" male, producing in F_1 the classes curly and wild in equal numbers. These F_1 flies were heterozygous for the "IIple" genes. An F_1 curly female was then crossed back to a "IIple" male, producing a back-cross generation which showed unexpected results. Instead of a general reassortment of characters of the "IIple" genes, there were no recombinations at all. What is probably

to be regarded as a 1:1 ratio of curly and "IIple" flies was obtained, which demonstrates very clearly the suppression of crossing over due to curly or to some gene or genes linked to it. Trial after trial of this same sort of mating was made under normal conditions, and there were no crossovers between curly and any of the "IIple" genes (table 4).

TABLE 4

Back-cross to test for the crossing over of curly and the "IIple" genes, using an F_1 curly female heterozygous for "IIple," and a "IIple" male.

DATE	CURLY $\frac{Cy}{\text{"IIple"}}$	"IIple"	CROSSOVERS	TOTAL
June 1921	151	74	0	225
July 1921	126	113	0	239
August 1921	368	257	0	625
September 1921	723	262	0	985
October 1921	299	114	0	413
Grand total	1667	820	0	2487

The ratio of the total number of "IIple" flies to that of the curly, designated in this table, gives an aberrant proportion on account, undoubtedly, of the low viability of the "IIple" flies in contrast with the extremely high viability of the curly flies. It has been noticed that where curly and wild flies are growing side by side in bottles in the same environment, if adverse conditions arise, the wild flies die before the curly. The "IIple" flies are much more susceptible to harmful external conditions than the wild flies, as is characteristic of mutants, and it is apparent that in a generation in which both curly and "IIple" flies are produced, many of the "IIple" flies might die and the curly would live.

In some of the bottles where conditions were good, as judged from the appearance of the food and the general vigor of the flies, a 1:1 ratio of curly and "IIple" flies was produced. All of these flies were raised in vials, following the method used by CLAUSEN. But CLAUSEN has recently shown, by comparative tests, that, unless these vials are fed with several times as much food as was customary, the ratios obtained are very poor. This is especially true where vestigial is involved, and is also true of "IIple" and "Xple." Unfortunately this defect of the vial method was not known at the time these experiments were made, and so the ratios depart rather widely from expectation in some instances. With this in mind, only conclusions that are unaffected, or are only affected in degree, by the aberrations, have been drawn from the data.

The important feature of this experiment is not the ratio of curly to "IIple" flies, but the fact that in a total of 2487 flies obtained from the curly "IIple" back-cross, not a single crossover occurred. The significance of the fact becomes greater when one considers the possibility of the large number of different combinations such a seven-point back-cross as this one presents. The curly gene, or more probably some gene or genes linked to it, possesses the power to affect the crossing over in the second chromosome.

It is later demonstrated in this paper that there are at least two such "little crossover" factors present in the curly chromosome, one governing crossing over in the left end of the chromosome (table 6) and one governing crossing over in the right end (table 5). Curly crossover flies which were finally obtained under high temperature had lost the "little crossover" factor from the right end of the chromosome, and when mated displayed about the usual amount of crossing over in that region. The curly crossover flies $\left(\frac{C_y\ p_r\ c\ p_z\ s_p}{+\ +\ +\ +\ +}\right)$, heterozygous for purple, curved, plexus, and speck, were mated to males of the same genetic constitution. Table 5 gives the number of crossovers obtained. Since crossing over takes place to a considerable extent in the right end of the second chromosome with curly when the original right end is missing, and does not take place under normal temperature conditions when the original right end is present, there is a factor in the right end that prevents crossing over. The factor may be called C_{CR}

TABLE 5

Offspring from two flies each heterozygous for curly, purple, curved, plexus, and speck. Here the curly chromosome contains the left end of the original curly chromosome and the right end of the "IIple" chromosome.

Parents: $\frac{C_y\ p_r\ c\ p_z\ s_p}{+\ +\ +\ +\ +}$ ♀ × $\frac{C_y\ p_r\ c\ p_z\ s_p}{+\ +\ +\ +\ +}$ ♂

No crossing over.....	Curly Wild	32 107
Crossing over between	**Visible characters**	**Number of flies**
2, p_r and c	$C_y\ c\ p_z\ s_p$	3
3, c and p_z	$C_y\ p_z\ s_p$	8
4, p_z and s_p	$C_y\ s_p$	4
2, 3	$C_y\ c$	2
3, 4	$C_y\ p_z$	2
1, 2, 4	$C_y\ p_r\ s_p$	4

To test for crossing over between star and curly, F_1 star curly females, having star in one second chromosome and curly in the other, were mated to wild males and put in a temperature of 30°C. No crossovers were obtained from 877 flies (table 6).

TABLE 6

Females heterozygous for star and curly in opposite chromosomes crossed to wild males and kept in a temperature of 30°C exhibit no crossing over.

NON-CROSSOVERS		CROSSOVERS	TOTAL
Curly	Star	Curly star	
448	429	0	877

The flies were reared at 30°C for a length of time necessary to affect the germ cells and to partly counteract the effect of the crossover factor C_{CR} in the right end of the curly chromosome. No crossing over occurred under these conditions between star and curly; therefore it appears that there is another crossover factor in the left end of the curly chromosome which is not affected by a temperature of 30°C and which prevents crossing over between star and curly. This crossover factor may be called C_{CL}.

Added proof of the presence of C_{CL} in the left end of the curly chromosome was obtained. Crosses were made in a temperature of 30°C, using curly crossover flies possessing the left end of the curly chromosome but not the right. The effect of the crossover factor C_{CR} in the right end of the curly chromosome was therefore eliminated. Curly purple curved plexus speck crossovers were mated to star flies, producing in F_1 some star curly flies. F_1 star curly females were crossed to wild males, producing classes star and curly but no crossovers, star curly, or wild. Of 518 flies, 264 were curly and 254 star.

EFFECT OF THE CROSSOVER FACTORS ON CHROMOSOMES I AND III

Since the crossover factors affect crossing over in the second chromosome, it was decided to test their behavior with regard to the first chromosome. It is known that crossover factors in one chromosome may modify the crossing over in other chromosomes. This is illustrated by the action of $C_{III\text{-}II}$, a gene in chromosome III that modifies crossing over in chromosome II (STURTEVANT 1919), and by $C_{II\text{-}III}$, a second-chromosome gene that modifies crossing over in the third chromosome (BRIDGES, unpublished).

Matings were made between curly females and males from the double recessive multiple stock of first-chromosome characters called "Xple." This stock contains the genes for scute bristles, locus at 0, echinus eye at 5.5, cut wing at 20, vermilion eye at 33, garnet eye at 44.5, and forked bristles at 56.5 (BRIDGES 1921). An F_1 curly female heterozygous in her first chromosome for the "Xple" genes, and in her second chromosome for curly and for the "little crossover" factors, was back-crossed to an "Xple" male to test for crossing over.

One hundred ninety-three flies were counted and the crossover classes were present in sufficient numbers to show that the crossover factors in

TABLE 7

Back-cross of F_1 curly females to "Xple" males.

	VISIBLE CHARACTERS	NUMBER OF FLIES
No crossing over	Curly "Xple"	12
	Curly	66
	"Xple"	7
	Wild	64
Crossing over between		
1, s_e and e_e	$C_y\ s_e$	2
	$C_y\ e_e\ c_t\ v\ g\ f$	1
2, e_e and c_t	s_e	1
	$C_y\ s_e\ e_e$	2
	$s_e\ e_e$	2
3, c_t and v	$C_y\ s_e\ e_e\ c_t$	1
	$v\ g\ f$	1
4, v and g		0
5, g and f	$C_y\ f$	3
	f	2
1, 2	e_e	3
1, 4	$C_y\ s_e\ g\ f$	1
1, 5	$C_y\ s_e\ f$	1
	$s_e\ f$	1
2, 3	$C_y\ c_t$	3
	c_t	3
2, 4	$c_t\ v$	1
2, 5	$s_e\ e_e\ f$	4
3, 4	$s_e\ e_e\ c_t\ g\ f$	1
3, 5	$s_e\ e_e\ c_t\ f$	1
	$v\ g$	2
4, 5	g	3
1, 2, 4	$e_e\ g\ f$	1
1, 2, 5	$e_e\ f$	2
2, 3, 4	$c_t\ g\ f$	1
2, 3, 5	$C_y\ c_t\ f$	1
Totals		193

the second chromosome have no great effect upon crossing over in the first chromosome (table 7).

The small number of flies in the expected classes, curly "Xple" and "Xple," recorded in table 7 is due to low viability of "Xple" flies in contrast to the curly and wild under the conditions of culture.

When a calculation of the percentages of crossing over is made from the data of table 7, it is found that the total amount of crossing over in section 1 is 6.7 percent, as compared with standard 5.5 percent; in section 2, 11.9 *versus* standard 14.5; in section 3, 7.3 *versus* 13.0; section 4, 4.2 *versus* 11.5; section 5, 10.9 *versus* 12.0. Of these values the only one significantly different from standard is 4 (4.2 *versus* 11.5) and probably 3 (7.3 *versus* 13.0). But when the data are looked at more closely, it is seen that this crossing over involves many times as much double and triple crossing over as in the standard condition. This fact, together with the differences in sections 4 and 3, means that the presence of curly and its associated modifiers has had a definite modifying effect upon crossing over in the X chromosome. The amount by which this effect differs from the standard is not so certain because of the small numbers raised and because of the marked disturbances to viability from the use of the vial method.

In a manner similar to that employed in testing the effect of the crossover genes on the first chromosome, experiments to test their behavior in regard to the third chromosome were performed. Matings were made, using curly females and males from the homozygous recessive multiple stock of third-chromosome characters called "IIIple." This stock contains the genes for roughoid eye, locus at 0, hairy body at 26.5, scarlet eye at 43.8, peach eye at 48.0, spineless bristles at 58.5, and sooty body at 70.7 (MORGAN, STURTEVANT and BRIDGES 1922, page 326).

An F_1 curly female heterozygous for the crossover factors and for the genes of "IIIple," was back-crossed to a "IIIple" male to test for crossing over. One hundred and fifty flies were counted, as is shown in table 8.

The percentages of crossing over calculated from the data of table 8, as compared with the standard, are given in table 9.

From this comparison it is apparent that there has been a very great increase in crossing over in section 3. The observed value is nearly six times the standard, and is the largest instead of the smallest of those involved. Section 1 showed only about half the standard amount of crossing over. The other values are nearly standard. From these data, and from the symmetry of the distribution of crossing-over effects observed by BRIDGES (1920) for the third chromosome, it appears that the crossing-over modifiers of curly increase crossing over in the middle of

TABLE 8

Back-cross of F_1 curly females to "IIIple" males.

	VISIBLE CHARACTERS	NUMBER OF FLIES
No crossing over	C_y "IIIple"	12
	C_y	21
	"IIIple"	12
	Wild	21
Crossing over between		
1, r_u and h	$C_y\ r_u$	6
	$C_y\ h\ s_t\ p^p\ s_s\ e^s$	3
	r_u	1
	$h\ s_t\ p^p\ s_s\ e^s$	1
2, h and s_t	$C_y\ s_t\ p^p\ s_s\ e^s$	1
	$C_y\ r_u\ h$	3
	$s_t\ p^p\ s_s\ e^s$	2
	$r_u\ h$	2
3, s_t and p^p	$C_y\ p^p\ s_s\ e^s$	7
	$C_y\ r_u\ h\ s_t$	4
	$p^p\ s_s\ e^s$	8
	$r_u\ h\ s_t$	3
4, p^p and s_s	$C_y\ s_s\ e^s$	5
	$C_y\ r_u\ h\ s_t\ p^p$	2
	$C_y\ e^s$	2
5, s_s and e^s	e^s	1
	$r_u\ h\ s_t\ p^p\ s_s$	1
1, 2,	$C_y\ r_u\ s_t\ p^p\ s_s\ e^s$	2
	$r_u\ s_t\ p^p\ s_s\ e^s$	1
1, 5,	$C_y\ r_u\ e^s$	2
2, 3,	$C_y\ r_u\ h\ p^p\ s_s\ e^s$	7
	$r_u\ h\ p^p\ s_s\ e^s$	3
2, 4,	$s_t\ p^p$	1
2, 5,	$C_y\ r_u\ h\ e^s$	2
	$r_u\ h\ e^s$	1
3, 4,	$C_y\ p^p$	1
3, 5,	$C_y\ r_u\ h\ s_t\ e^s$	2
	$r_u\ h\ s_t\ e^s$	1
	$p^p\ s_s$	1
4, 5,	$C_y\ s_s$	1
	$r_u\ h\ s_t\ p^p\ e^s$	1
	s_s	1
1, 4, 5,	$r_u\ s_s$	1
2, 4, 5,	$C_y\ s_t\ p^p\ e^s$	1
	$s_t\ p^p\ e^s$	1
3, 4, 5,	$C_y\ p^p\ e^s$	1
	$p^p\ e^s$	1
Total		150

the chromosome very greatly, that this effect falls to normal not far from the middle on each side, and then falls below normal in the ends. In these data there is the same large disproportion of double and of triple crossovers as is the case in chromosomes I and II. These results are also questionable from the small numbers obtained, and from the use of the vial method. Another source of error not adequately controlled was the

TABLE 9

Crossing-over section	1	2	3	4	5
Values from table 8	11.3	18.0	26.0	11.3	14.0
Standard values	26.5	17.0	4.2	10.0	12.5

age of the mothers. BRIDGES finds (unpublished data) that the first output of eggs of females heterozygous for third-chromosome genes shows an abnormally high amount of crossing over in section 3, and that in the offspring of very old females, this may rise as high as the percentage observed here. However, it is believed that the deviations of table 9 do show an effect upon crossing over in chromosome III by the modifiers associated with curly.

INCREASE OF CROSSING OVER DUE TO A HIGHER TEMPERATURE

No crossing over occurred between curly and any of the "IIple" genes in a curly "IIple" back-cross until July, 1921, when the flies were kept in an incubator at a high temperature at Woods Hole. It was not possible to keep the degree of temperature constant on account of lack of equipment, but a temperature of from 25°C to 30°C was attempted. After two generations in the incubator, crossovers were found between black and purple, and purple and curved. Flies from bottles in the incubator were taken as parents for the same back-cross and were carried around in a traveling bag for four weeks. From these bottles no crossover flies appeared. It looked as though temperature were the factor responsible for the crossing over, and an experiment was commenced involving temperature differences.

F_1 curly females heterozygous for the "IIple" characters were again used in a back-cross and were placed in the incubator in a temperature of 26°C. Bottles were likewise placed outside the incubator in ordinary room temperature of about 20°C for a control. No crossovers were produced in any of the bottles either inside or outside of the incubator (table 10).

TABLE 10

Back-cross of F_1 curly females by "IIple" males.

TEMPERATURE	CURLY	"IIPLE"	CROSSOVERS	TOTAL
20°C	335	242	0	577
26°C	203	25	0	228

A new experiment was started and the temperature of the incubator raised to 30°C. The most favorable degree of heat for normal *Drosophila melanogaster* is said to be about 25°C. In working with curly flies, I find that they thrive in 30°C and show an increase of viability in that temperature. Table 11 gives numbers of flies of various combinations of characters hatched at 30°C, and table 12 shows the percentages of crossing over as calculated from table 11. Even at the high temperature the crossing over is not as frequent as it is in the same regions in other stocks than curly. A total of 1075 flies from this curly × "IIple" back-cross produced 38 crossovers. Of these an abnormally high proportion were double and triple crossovers. The greatest proportional increase in crossing over took place between black and purple, and between plexus and speck. Crossing over between curved and plexus was least increased by the high temperature.

TABLE 11

Back-cross of F_1 curly females by "IIple" males at 30°C.

	VISIBLE CHARACTERS	NUMBER OF FLIES
No crossing over	"IIple"	409
	Curly	628
Crossing over between		
2, *b* and p_r	$C_y\ p_r\ c\ p_z\ s_p$	7
	$T^{dy}\ b$	2
3, p_r and *c*	$C_y\ c\ p_z\ s_p$	6
	$T^{dy}\ b\ p_r$	2
4, *c* and p_z	$C_y\ p_z\ s_p$	1
5, p_z and s_p	$C_y\ s_p$	3
2, 3	$C_y\ p_r$	6
2, 5	$T^{dy}\ b\ s_p$	1
3, 4,	$C_y\ c$	2
2, 3, 5,	$C_y\ p_r\ s_p$	1
2, 4, 5,	$C_y\ p_r\ c\ s_p$	4
3, 4, 5,	$C_y\ c\ s_p$	3
Total		1075

TABLE 12

Percentages of crossing over in each of the regions of the second chromosome tested in the experiment shown in table 11.

CROSSING OVER BETWEEN	TOTAL CROSSOVERS	PERCENT	STANDARD PERCENT
2, *b* and p_r	21	1.95	6.0
3, p_r and *c*	20	1.86	21.0
4, *c* and p_z	10	0.93	25.0
5, p_z and s_p	12	1.11	6.5
Total		5.86	58.5

CROSSOVER FACTORS IN THE HOMOLOGUE OF THE CURLY CHROMOSOME

In an attempt to locate a lethal linked to the gene for vestigial in the homologue of the curly chromosome, the presence of a "little crossover" factor was discovered in that chromosome. From a P_1 mating of a star male and a curly female, F_1 star not-curly females were used for back-crosses to curly males. Only the non-crossover classes, star curly, star, and curly, were present. No wild-type, star vestigial, or vestigial flies appeared. Therefore we may justifiably assume that no crossing over occurred (table 13). Some gene or genes located in the chromosome carrying the gene for vestigial is responsible for the non-crossing over of genes within the second chromosome. From the early history of the curly stock it seems probable that the curly mutation arose in a stock that was already homozygous for these crossing-over modifiers, and that the modifiers present in the vestigial-bearing homologue are the same as those present in the curly homologue. This becomes practically a certainty when the effect of temperature upon the crossing-over modifiers of the not-curly homologue are compared with those of the curly homologue.

TABLE 13

Cross of star females heterozygous for the homologue of the curly chromosome, and curly males, in a temperature of 20°C.

STAR CURLY	STAR	CURLY	TOTAL
132	66	175	373

EFFECTS OF TEMPERATURE ON THE CROSSOVER FACTORS IN THE HOMOLOGUE OF THE CURLY CHROMOSOME

To test for the effect of temperature on the crossover factors in the homologue of the curly chromosome, F_1 star females heterozygous for the

homologue of the curly chromosome were back-crossed to curly males and placed in a temperature of 30°C. All the non-crossover classes, star, star curly, and curly, were produced, and the crossover classes, star vestigial, and wild. Table 14 gives the results of this cross. Of a total of 121 flies, 22 were crossovers besides those star curly and curly flies which are included in the non-crossover classes in the table. Since there was no crossing over under normal temperature conditions, and some crossing over in the high temperature, the high temperature is the factor responsible for the partial counteraction of the effect of the crossover factor in the homologue of the curly chromosome. This behavior is similar to that found for the crossing-over modifiers of the curly homologue and shows that the two sets are most probably the same.

The cross of table 14 shows, by the absence of the vestigial class, that the order of the genes is star, lethal-vestigial, vestigial. Furthermore, since the data of table 12 shows that there is only negligible crossing over to the left of black, the locus of the lethal is about half-way between the loci for black and for vestigial, or at about 10 units to the left of vestigial.

TABLE 14

Cross of star females heterozygous for the homologue of the curly chromosome, and curly males, in a temperature of 30°C.

	VISIBLE CHARACTERS	NUMBER OF FLIES	
No crossing over.....	$S\ C_y$ Curly Star	41 42 16	These include crossovers
Crossing over between 1, S and l_{vg} 2, l_{vg} and v_g 1, 2,	 Wild $S\ v_g$ v_g	 10 12 0	
Total		121	

LOCUS OF THE GENE FOR CURLY

The 38 crossover flies obtained from the back-cross between curly and "IIple" at 30°C (see table 12) showed no crossing over between curly and black. No curly black fly appeared in my experiments, but Professor CLAUSEN reports one curly black and two curly black purple flies in a total of 1325 flies of a curly "IIple" back-cross generation. These data

show that curly is to the left of black. If curly were located to the right of black, black curly purple double crossovers could hardly be obtained, for the distance is too short. If it were not for the influence of the "little-crossover" factor in the left end of the curly chromosome, which prevents crossing over of genes in that end, more data might be obtained to locate the exact position of curly in the second chromosome.

CINNABAR[2], A CINNABAR ALLELOMORPH IN THE CURLY CHROMOSOME

To show that the curly and the vestigial genes of the curly stock were in different homologues, matings were made between F_1 star curly dichaete flies, from a cross of curly by star dichaete, and cinnabar vestigial. Cinnabar, a bright vermilion eye color, is a new second-chromosome recessive located to the right of purple at 55.5. It was found by Professor CLAUSEN, but he has not yet published his results. If curly and vestigial separate at reduction, such a cross of an F_1 star curly dichaete fly and a cinnabar vestigial fly should give an F_2 that contains no vestigial flies. This was the case; no vestigial flies appeared, but an unexpected new feature appeared, in that all the curly flies had cinnabar-like eyes. Cinnabar being a recessive character, this peculiar phenomenon could not have taken place unless the curly chromosome has a deficiency for cinnabar, or carries a gene for cinnabar, or an allelomorph of cinnabar.

The color of the cinnabar eye is very similar to vermilion, and these curly cinnabar-like flies were not quite as bright in color as flies homozygous for cinnabar. On account of this difference in color, the existence of a cinnabar-deficiency in the curly chromosome might have been doubted from the first. However, since the "deficiency" idea fitted in well with the lethal effect of the curly chromosome when homozygous, further evidence was obtained by testing for deficiency.

Cinnabar is located at 55.5, purple at 53, and safranin at 58. If the curly chromosome is deficient for cinnabar, it might be deficient also for purple or safranin or for both. Curly flies were crossed to homozygous purple and to homozygous safranin flies, but in neither case did any of the F_1 curly flies show the eye color of the homozygous parent. No deficiency for purple or safranin, therefore, exists in the curly chromosome.

The dumpy black crossover flies obtained from a curly × "IIple" back-cross at high temperature, contains the right end of the curly chromosome in one of the second chromosomes:

$$\frac{T^{dy}\ b\ \text{(right end of curly chromosome)}}{T^{dy}\ b\ p_r\ c\ p_z\ s_p}$$

If no deficiency exists in that chromosome, it must possess a gene for cinnabar, or its allelomorph, in heterozygous form. A simple cross of two such flies will serve to demonstrate whether or not such a gene is really present in the curly chromosome. Crossovers, $\left(\frac{T^{dy}b\ c_n{}^2?}{\text{"IIple"}}\ ♀ \times \frac{T^{dy}b\ c_n{}^2?}{\text{"IIple"}}\ ♂\right)$, mated together, produced 55 "IIple," 30 dumpy black cinnabar flies, and 40 dumpy black flies. The presence of cinnabar-like flies (homozygous) showed that there is no lethal factor in the right end of the curly chromosome, and since there is no lethal effect there is no deficiency.

The color of these homozygous cinnabar-like flies was not nearly as bright as that of Professor CLAUSEN's cinnabar, and was not as bright as the cinnabar of the curly flies of the back-cross discussed above. This was taken to prove that the cinnabar gene in the curly chromosome is not the same as the original cinnabar but is an allelomorph that produces less effect. The intermediate color of the eyes of the curly cinnabars is due to the fact that they are "compounds" between the bright cinnabar and the dull allelomorph. A further test of the character of this cinnabar allelomorph, cinnabar^2, was made by mating heterozygous dumpy black cinnabar^2 flies to cinnabar. One-half of the offspring had cinnabar eyes (c_n–$c_n{}^2$ compounds), that were not quite as bright as CLAUSEN's cinnabar, and the other half had wild-type eyes. Homozygous dumpy black cinnabar^2 flies were mated to cinnabar, and all of the offspring had cinnabar eyes not quite as bright as the cinnabar parent.

EVALUATION OF CURLY

Curly is a dominant heterozygous character which is easily recognizable. These facts make it valuable for identification purposes in genetic work on *Drosophila melanogaster*. In normal temperature the "little-crossover" factors contained in the curly chromosome restrict all crossing over along the entire length of the second chromosome, and keep genes in the two chromosomes from crossing over with each other. When prevention of crossing over is desirable, curly flies may be used. No microscope is necessary in separating curly-winged flies from others, as the upturned wings attract attention at once. Curly flies possess a very high viability and can live in conditions adverse to the health of the wild fly. They are indeed much hardier than any other stocks that I have observed. A pair of curly flies produces a large number of offspring and since practically all flies homozygous for the curly chromosome die, those that do hatch represent a large percent of the number of viable eggs laid. Curly-

winged flies are more easily handled since they are able to fly only in short jumps. They respond quickly to light and are easily transferable from one bottle to another. The genetic and physiological qualities of the curly flies make them especially desirable for experimental work with *Drosophila melanogaster*.

SUMMARY

1. Curly is a definite easily classifiable dominant mutant wing character.

2. The gene for curly is located in the second chromosome. It is usually lethal in homozygous form. The homozygous curly type is easily separable from the ordinary heterozygous type.

3. The homologue of the curly chromosome in the curly stock contains a gene for vestigial and also a lethal. This lethal gives with curly a pure-breeding balanced stock of curly. This balanced stock arose during the course of selection, and is similar in its genetic behavior to beaded, truncate, and double-flowered Matthiola. Twin hybrids are produced when curly is crossed to wild stock.

4. Two "little-crossover" factors, one in the left end of the curly chromosome, C_{CL}, and one in the right end, C_{CR}, prevent crossing over in the second chromosome under normal temperature conditions. The effect of the crossover factor in the right end is reduced when the temperature is raised to 30°C, but the one in the left end is not affected by a rise in temperature.

5. The crossover factors in the curly chromosome probably affect crossing over in the first chromosome, and also in the third chromosome.

6. In the vestigial-bearing homologue crossover factors exist that are probably the same as those in the curly homologue.

7. At 30°C curly shows crossing over with black only very rarely, more frequently with purple and with curved. The crossover relations indicate its locus to be to the left of black.

8. The curly chromosome contains a gene for an eye color similar to vermilion but less bright, and allelomorphic to cinnabar; it is designated cinnabar^2 (c_n^2).

9. The curly second chromosome is valuable genetically on account of the dominance of its curly factor, on account of its low-crossover factors, and its recessive lethal effects.

10. Vestigial and cinnabar^2 flies occasionally appear in the pure curly stock, simulating mutation, although they are really due to crossing over of vestigial and cinnabar from the balancing lethals.

11. The genetic constitution of all flies in the pure-breeding curly stock may be represented as $\frac{C_{CL}\, C_v + c_n^2 + C_{CR}}{C_{CL} + l_w + vg^2\, C_{CR}}$.

LITERATURE CITED

ALTENBURG, EDGAR, and MULLER, HERMANN J., 1920 The genetic basis of truncate wing, an inconstant and modifiable character in Drosophila. Genetics **5**: 1–59.

BRIDGES, CALVIN B., 1921 Current maps of the location of the mutant genes of *Drosophila melanogaster*. Proc. Nation. Acad. Sci. **7**: 127–132.

BRIDGES, C. B., and MORGAN, T. H., 1919 The second-chromosome group of mutant characters. Carnegie Inst. Washington Publ. 278, 388 pp.

DEXTER, J. S., 1914 The analysis of a case of continuous variation in Drosophila by a study of its linkage relations. Amer. Nat. **48**: 712–758.

FROST, H. B., 1915 The inheritance of doubleness in Matthiola and Petunia. I. The hypothesis. Amer. Nat. **49**: 623–636.

1919 Mutation in Matthiola. Univ. of California Publ. **2**: 81–190.

MORGAN, T. H., STURTEVANT, A. H., and BRIDGES, C. B., 1922 Study of the constitution of the germ-plasm in relation to heredity. Carnegie Inst. Washington Year-book **21**: 325–328.

MULLER, HERMANN J., 1918 Genetic variability, twin hybrids and constant hybrids, in a case of balanced lethal factors. Genetics **3**: 422–499.

PLOUGH, HAROLD H., 1917 The effect of temperature on crossing over in Drosophila. Jour. Exp. Zoöl. **24**: 147–209.

SAUNDERS, E. R., 1911 Further experiments on the inheritance of doubleness and other characters in stocks. Jour. Genetics **1**: 303–376.

1913 On the mode of inheritance of certain characters in double-throwing stocks. A reply. Zeitschr. indukt. Abstamm. u. Vererb. **10**: 297–310.

VRIES, H. DE, 1918 Twin hybrids of *Oenothera Hookeri* T and G. Genetics **3**: 397–421.

JOURNAL *of* GENETICS

Edited by

W. BATESON AND R. C. PUNNETT

Founded in 1910, the Journal of Genetics is the only British periodical devoted to the publication of original research in Heredity and Variation. Up to the end of 1922 twelve volumes have been completed. The illustrations form a feature of the series, for, in addition to numerous text figures and diagrams, the twelve volumes contain 262 plates, of which 54 are in colours.

Many of the papers published deal with animals and plants of high economic value, and are therefore of interest to those concerned with agriculture and horticulture.

A full index to Vols. I-XII was issued with the last part of Vol. XII.

The Journal of Genetics is published in parts, of which three form a volume. Approximately one volume is issued annually. The subscription price for a volume is £2:2:0 net; the price of the separate parts being shillings 15 net each. The parts are sent post free to subscribers as issued.

Subscriptions, payable in advance, should be sent to Mr. C. F. Clay, Cambridge University Press, Fetter Lane, London.

The Cambridge University Press has appointed THE UNIVERSITY OF CHICAGO PRESS agent for the sale of the Journal of Genetics in the United States of America, and has authorized the following prices:—Annual subscription $10.00 net; single copies $3.50 net each.

Inquiry as to back numbers should be made of Mr. C. F. Clay. Separate parts and volumes are still available in some cases.

Brooklyn Botanic Garden Memoirs

Volume II

The Vegetation of Long Island

Part I

The Vegetation of Montauk

A Study of Grassland and Forest

By NORMAN TAYLOR

Price, $1.00. Remittance should accompany all orders. Published June 11, 1923. Not offered in exchange.
108 pages, 30 text figures

Address

The Secretary, BROOKLYN BOTANIC GARDEN, *Brooklyn, N. Y., U. S. A.*

Ecology

Devoted to All Forms of Life in Relation to Environment

Established 1920 Quarterly

Official Publication of the Ecological Society of America

Subscription, $4 a year. Back volumes I-III, $3.50 each, post free
Foreign postage: To Canada 10 cents; to other countries, 20 cents.
Single numbers $1.15 post free.

Published by BROOKLYN BOTANIC GARDEN, *Brooklyn, N. Y., U. S. A.*

Bibliographic Service in BOOK FORM, Volume 1

1922. Bibliographic Service in book form with authors' and analytical subject indices, covering two and a half years of the Bibliographic Card Service, June 1, 1917, to December 30, 1919, cards nos. 1 to 519 **Price, $5.00**

To subscribers to the Bibliographic Service, covering this period, 50 per cent discount

Volume II in preparation

THE WISTAR INSTITUTE OF ANATOMY AND BIOLOGY
PHILADELPHIA

INFORMATION FOR CONTRIBUTORS

The length of contributions to GENETICS will not be arbitrarily limited, but articles of less than fifty pages may be expected to appear more promptly than those of greater length. The limitations as to subject-matter will be construed broadly, so that any paper of sufficient importance, whether in the field of genetics proper, of cytology, embryology, physiology, biometry, or mathematics, if of such character as to be of primary interest to the geneticist, will find here an appropriate place of publication.

Contributors are requested to observe the usual care in the preparation of manuscripts. All references to literature should cite the name of the author, followed by the year of publication, the papers so referred to being collected into a list of "Literature cited" at the end of the article. In this list great care should be taken to give titles *in full*, and to indicate accurately in Arabic numerals, the volume number, and first and last pages, and the date of publication of each paper, if published in a periodical, and the number of pages, place and date of publication and name of publisher, of each independent publication. The arrangement of this list should be alphabetical by authors, and chronological under each author.

Footnotes, legends for figures, tabular matter, and quotations involving more than one sentence should be typewritten on separate sheets. All footnotes should be numbered consecutively in a single series and designated by Arabic superscript numerals.

No limitation will be placed on the kind of illustration to be used, but as funds for illustration are still very limited, authors are requested to practice restraint, and to limit themselves to such amount and grade of illustration as seem positively essential to an adequate understanding of their work. The size of the type-forms of GENETICS, exclusive of the heading, is $4\frac{5}{8} \times 7$ inches and all illustrations intended for use as single plates should not exceed these dimensions or should allow of reduction to them. Double plates may be $12\frac{1}{2} \times 7$ inches.

Authors will receive galley-proofs, and, if specially requested, page-proofs will be sent also, it being understood that the latter must be returned very promptly, and that no extensive change may be made in the page-proofs, which is not compensated for within the same paragraph, or in an adjacent paragraph on the same page. Changes from copy will be charged against the author.

Contributors will receive 50 reprints of their papers, with covers, free of charge, provided not fewer than 100 are ordered. If a less number is desired, one-half the total cost of their manufacture will be charged to the author. Additional copies may be secured at actual cost of manufacture, provided these are ordered when the corrected galley-proofs are returned.

Manuscripts and all editorial correspondence should be addressed to the Managing Editor, DR. GEORGE H. SHULL, 60 Jefferson Road, Princeton, New Jersey, U.S.A.

GENETICS, MAY 1923

TABLE OF CONTENTS

PAGE

SAX, KARL, and GOWEN, JOHN W., Permanence of tree performance in a clonal variety, and a critique of the theory of bud mutation 179

COLLINS, J. L., and MANN, MARGARET C., Interspecific hybrids in Crepis. II. A preliminary report on the results of hybridizing *Crepis setosa* Hall with *C. capillaris* (L.) Wallr. and with *C. biennis* L. 212

DE VRIES, HUGO, and BOEDIJN, K., On the distribution of mutant characters among the chromosomes of *Oenothera Lamarckiana* 233

WRIGHT, SEWALL, The theory of path coefficients—a reply to Niles's criticism 239

NILES, HENRY E., The method of path coefficients—an answer to Wright 256

BOSHNAKIAN, SARKIS, The relation of the spelt factor in wheat to rachis internode characters 261

WARD, LENORE, The genetics of curly wing in Drosophila, another case of balanced lethal factors 276